NAVAL RESEARCH LOGISTICS QUARTERLY

DECEMBER 1975
VOL. 22, NO. 4

OFFICE OF NAVAL RESEARCH

NAVSO P-1278

NAVAL RESEARCH LOGISTICS QUARTERLY

The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.

The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $10.00 a year in the U.S. and Canada, $12.50 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.

Issuance of this periodical approved in accordance with Department of the Navy Publications and Printing Regulations, P-35 (Revised 1-74).

ON THE TREATMENT OF FORCE-LEVEL CONSTRAINTS IN TIME-SEQUENTIAL COMBAT PROBLEMS*

James G. Taylor

Naval Postgraduate School
Monterey, California

ABSTRACT

The treatment of force-level constraints in time-sequential combat optimization problems is illustrated by further studying the fire-programming problem of Isbell and Marlow. By using the theory of state variable inequality constraints from modern optimal control theory, sharper results are obtained on necessary conditions of optimality for an optimal fire-distribution policy (in several cases justifying conjectures made in previous analysis). This leads to simplification of the determination of the domains of controllability for extremals leading to the various terminal states of combat. (Additionally, some new results for the determination of boundary conditions for the adjoint variables in optimal control problems with state variable inequality constraints have arisen from this work.) Some further extensions of previous analysis of the fire-programming problem are also given. These clarify some key points in the solution synthesis. Some important military principles for target selection and the valuation of combat resources are deduced from the solution. As a result of this work, more general time-sequential combat optimization problems can be handled, and a more systematic solution procedure is developed.

1. INTRODUCTION

The determination of target priority rules for fire distribution/target selection frequently arises in defense planning studies such as the evaluation of proposed weapon systems or force mixes (see, for example, [20] for a discussion within the context of the evaluation of fire-support systems). In particular, the determination of optimal time-sequential fire-distribution strategies for supporting weapon systems † is a major problem of military operations research. Early work was done on this problem at RAND in the late 1940's and early 1950's (see [8]) and elsewhere (see [2]). Today the problem of determining optimal time-sequential air-war strategies is being studied by a number of organizations (see, for for example [9], [16], [26], [36]). Such problems were extensively discussed in the workshop on optimization techniques and combat applications at the (1973) Conference on the State-of-the-Art of Mathematics in Combat Models (see [24]; also [34]) and in working groups at the 31st and 32nd Military Operations Research Symposia (also held in 1973). Considering such significant applications, it is important to have both mathematical solution methodology available for such problems and also a complete mathematical treatment of the simplest time-sequential fire-distribution problem.

The time-sequential fire-distribution problem first studied by Isbell and Marlow (see [28]) is probably the simplest optimal control problem that arises in the Lanchester theory of combat (see [30]).

---

*This research was supported by the Office of Naval Research as part of the Foundation Research Program at the Naval Postgraduate School and partially under Project Order No. 2-0150.

† See [37] for a discussion of the distinction between a "primary" weapon system (or infantry) and a "supporting" weapon system.

Consequently, the development of a complete solution to this problem (along with the appropriate solution methodology) is significant for guiding extensions to more complex combat dynamics in both one-sided and also two-sided optimization problems. Such a complete solution is important not only for its own sake,* but also for testing the capabilities of proposed computational methods for time-sequential combat allocation problems (for example, Lagrange dynamic programming [25] for discrete-time versions of such problems (see also [16], [36])). The author [28] recently extended the pioneering 1956 work [14] of Isbell and Marlow † by determining domains of controllability for extremals** leading to the various terminal states of combat and discovering the dispersal surfaces present in the solution. We had to base, however, part of our previous analysis on several conjectures regarding boundary conditions for the adjoint variables. Additionally, subsequent research has revised some of our arguments regarding the determination of global optimality of extremal fire-distribution policies.

Thus, the purposes of the paper at hand are to extend some of our previous analysis on considerations "in the large" for the Isbell-Marlow fire-programming problem [28] and to introduce a more comprehensive approach (the theory of state variable inequality constraints (SVIC's)) for the treatment of force-level constraints in time-sequential combat optimization problems which yields sharper optimality results "in the small." The results of this further examination of the Isbell-Marlow problem are of considerable interest, since they provide an understanding of solution phenomena (multiple extremals and a dominated payoff) that we have previously encountered in a differential game [29]. Additionally, some important insights into military principles and the valuation of combat resources are deduced from the solution.

In [28] we showed that extremals for the Isbell-Marlow problem were not unique for a certain range of model parameters and tried to develop necessary and sufficient conditions for optimal paths to lead to all the terminal states of combat. In the pioneering 1956 work of Isbell and Marlow [14], one can find several gaps (boundary conditions for the adjoint variables, treatment of state variable inequality constraints (SVIC's)) in the light of subsequent control theory developments. In [28] we followed Isbell and Marlow's heuristic treatment of the determination of boundary conditions for the dual variables and did not consider the theory of SVIC's. We rectify these gaps here.

All the Lanchester-type (time-sequential) tactical allocation problems that the author has studied contain negativity restrictions on the force levels due to physical reasons (negative force levels do not make sense). In such models, force levels will be *state variables* so that such a restriction is mathematically called a state variable inequality constraint (SVIC). The maximum principle (in its original form [23]) is inadequate to handle SVIC's and special mathematical theories (see [15] for an almost comprehensive review) have been developed for such problems. When we started our research, we found that results were widely scattered in the literature and that a completely adequate theory to solve even the simplest fire-distribution problem did not exist. Hence, we have developed a theory of SVIC's for tactical allocation problems within the Lanchester theory of combat. The author would like to point out that his research on the "simplest" time-sequential fire-distributio

*One can clearly see the dependence of the structure of the optimal time-sequential fire-distribution policy on model form and model parameters.

† The Pontryagin maximum principle was only announced in 1956 [3], although Professor Magnus R. Hestenes of the United States had apparently first given the now standard control formulation (as well as developed optimality conditions by a translation of well-known classical calculus of variations results) and urged others to follow his approach in 1949 [10].

**By an extremal we mean a path (or trajectory) on which the necessary conditions of optimality are satisfied at every point in time. Thus, the optimal trajectory must be an extremal.

problem using the theory of SVIC's has led to some contributions [32], [33] to the control theory literature. This indicates the subtle difficulty of the Isbell-Marlow problem and the significance of our developments in the paper at hand.

In [31] we made the first application of the theory of SVIC's to allocation problems in the Lanchester theory of combat. There we applied the approach of Gamkrelidze (see ch. 6 in [23]) (also used by Bryson, Denham, and Dreyfus [5] (see also [15])). In the paper at hand we use the method of Speyer and Bryson* [27] of adjoining the SVIC directly to the criterion functional. This approach has several advantages, one of which is a richer military interpretation of the multiplier associated with a force level constraint. Moreover, our work here has led to an extension (see [33]) of Speyer's results [27].

Furthermore, we feel that one of our major contributions in [28] was the development of a *purely algebraic method* for the determination of the optimal control, as opposed to the *geometric method* employed in [14]. The reader should note the difficulty of applying this geometric approach to a similar problem with a state space of dimension higher than three. We feel that this is essentially impossible, whereas our algebraic method (like algebraic geometry in n-dimensional space) appears to be readily applicable to such problems.

2. THE ISBELL-MARLOW FIRE PROGRAMMING PROBLEM

For the reader's convenience we re-state the problem originally studied by Isbell and Marlow [14].

(1) $$\underset{\phi(t)}{\text{maximize}} \{ry(T) - px_1(T) - qx_2(T)\} \text{ with } T \text{ unspecified,}$$

subject to

$$\frac{dx_1}{dt} = -\phi a_1 y,$$

$$\frac{dx_2}{dt} = -(1-\phi) a_2 y,$$

$$\frac{dy}{dt} = -b_1 x_1 - b_2 x_2,$$

$$x_1, x_2, y \geq 0 \quad \text{and} \quad 0 \leq \phi \leq 1,$$

where all symbols are defined in the next section. In this problem x_1, x_2, and y are *state* variables, while ϕ is a *control* (or decision) variable. One contribution of this paper is to illustrate the mathematical treatment of a force level constraint, such as $x_1 \geq 0$.

---

*The author wishes to thank a referee for pointing out that the results [5], [15], and [27] have been formally deduced and for calling his attention to the rigorous developments [21], [22] of the late Professor Lucien Neustadt. Unfortunately, Neustadt's results for *measurable* controls cannot be directly applied to problems (such as the one at hand) with *piecewise continuous* controls, but require a translation. Moreover, Neustadt mentions neither the concept of a pth order SVIC (see [15]) nor that of an "absorbing" state boundary (see [33]).

The battle terminates upon reaching the terminal states defined by (1) $x_1(T) = x_2(T) = 0$ and (2) $y(T) = 0$. Upon further analysis, it has been convenient to consider that there are the following five "target sets" for this problem:

C_1: $x_1(T) = 0$, $x_2(T) > 0$, $y(T) = 0$,
C_2: $x_1(T) = 0$ before $x_2(T) = 0$, $y(T) > 0$,
C_3: $x_1(T) = 0$ after $x_2(T) = 0$, $y(T) > 0$,
C_4: $x_1(T) > 0$, $x_2(T) = 0$, $y(T) = 0$,
C_5: $x_1(T) > 0$, $x_2(T) > 0$, $y(T) = 0$.

The reader should note that in the above problem statement T is referred to as being undetermined. This is because T is determined by entry to one of the above five target sets. This is a function of the control applied, and hence before an allocation rule is given, it is unspecified.

3. NOTATION

The symbols which are used in this paper are defined as follows:

$A = A(R, z) = [z^2(R-1) - R]/(z-1)^2$,
$B = B(R, z) = A(z-1)^2/z^2 = [z^2(R-1) - R]/z^2$,
a_1, a_2, b_1, b_2 = constant attrition-rate coefficients,
C_i for $i = 1, 2, 3, 4, 5$ = the ith part of the terminal surface as defined in Section 2,
$D(C_i)$ = domain of controllability for C_i,
$g(P^0, R, z)$ = term in equation of the locus of points for which $P_1 = P_4$,
$h(P^0, R, z)$ = term in equation for boundary surface between the regions from which optimal paths lead to C_1 and C_4,
H = Hamiltonian function,
p, q, r = utilities assigned to surviving X_1, X_2, and Y forces, respectively,
$p_i(t)$ for $i = 1, 2, 3$ = dual variable corresponding to $x_i(t)$ $(x_3(t) = y(t))$,
P_i for $i = 1, 2, 3, 4, 5$ = payoff associated with an extremal leading to C_i,
$P^0 = (x_1^0, x_2^0, y_0)$ = point in the initial state space,
$Q = (-p_2(t=T))/q$,
$R = a_1b_1/(a_2b_2)$,
$s = s(x_1^0, x_2^0) = b_1x_1^0 + b_2x_2^0$,
t_1 = time at which X_1 is annihilated, i.e., $x_1(t_1) = 0$,
t_2 = first time at which $2b_1x_1(t_2)x_2^0 + b_2(x_2^0)^2 = a_2\gamma^2$ (t_2) for an extremal leading to C_4,
T = total time for the battle,
$v = v(\tau) = a_2p_2(\tau) - a_1p_1(\tau)$,

$$w = \cosh \sqrt{a_2b_2}\,\tau_1(C_4) = \frac{a_1}{p_2(t=T)} \frac{(b_1p_2(t=T) + b_2p)}{(a_1b_1 - a_2b_2)},$$

x_1, x_2, y = force levels; with initial values x_1^0, x_2^0, y_0,

$$z = \cosh \sqrt{a_2b_2}\,\tau_1(C_5) = \frac{a_1}{q} \frac{(b_1q - b_2p)}{(a_1b_1 - a_2b_2)} = \frac{R-\delta}{R-1},$$

$\delta = a_1 p/(a_2 q)$,

$\eta_i(t)$ for $i = 1, 2$ = multiplier corresponding to state variable inequality constraint $x_i \geq 0$,

ν_i for $i = 1, 2$ = multiplier corresponding to state variable terminal inequality constraint $x_i(T) \geq 0$,

ϕ = fraction of Y-fire directed at X_1,

τ = "backwards time" from the end of the battle defined by $\tau = T - t$, i.e., the time remaining before the end of the battle,

$\tau_1(C_i)$ = "backwards time" of the first switch in tactics for extremals leading to C_i.

ldditionally:

C_5^s refers to C_5 (the 5th part of the terminal surface) which is reached by extremals with a *switch* in tactics,

P_5^s = payoff associated with an extremal leading to C_5^s.

SUMMARY OF SOLUTION

Before we present our methodological advance of using the theory of SVIC's, let us consider e end result of our further analysis. It should be pointed out, however, that we have still followed e general solution procedure given in [28]. There are four cases to be considered:

(1) $\delta \geq 1$,
(2) $R - \sqrt{R(R-1)} < \delta < 1$,
(3) $\delta = R - \sqrt{R(R-1)}$,
(4) $0 \leq \delta < R - \sqrt{R(R-1)}$,

ere $\delta = a_1 p/(a_2 q)$. For Case (1): $\delta \geq 1$, the solution is given in Table 1 of [28]. Our re-examination re changes no details of the analysis presented there.

The results of our additional analysis (which includes use of the theory of SVIC's) are presented Tables I through VI. Supporting details will be presented in subsequent sections only when they ffer significantly from our previous analysis in [28].

For Case (2): $R - \sqrt{R(R-1)} < \delta < 1$, extremals are shown (in the sense that a battle trajectory rresponding to the extremal control leads from a point of the initial state space to the terminal rface) in Table I. By the domain of controllability for extremals leading to a given terminal state mean the subset of the initial state space such that a field of extremals leads to the terminal state. ie results given in Table I differ from those given in [28] in two important respects: the domains of ntrollability for extremals leading to C_1 and C_4, denoted as $D(C_1)$ and $D(C_4)$, have been revised. the case of $D(C_1)$ our previous analysis [28] had used unsupported and unstated assumptions that e not true in general. In the gase of $D(C_4)$ we have used some recent results [33] from the theory SVIC's about the boundary conditions for the dual (or adjoint) variables to obtain this new result.

It should be noted that $D(C_1)$ overlaps $D(C_4)$, $D(C_5)$, and $D(C_5^s)$. (However, $D(C_4)$, $D(C_5)$, d $D(C_5^s)$ are disjoint sets, i.e., they do not overlap each other.) Hence, except for this region of erlap the extremal control turns out to be the optimal control.* In the region of overlap further

*The existence of an optimal control is discussed in Section 7.

TABLE I. *Extremals for Isbell-Marlow Problem for* $R - \sqrt{R(R-1)} < \delta < 1$

Nonrestrictive Assumption: $R > 1$, i.e., $a_1 b_1 > a_2 b_2$.

Case (2): $R - \sqrt{R(R-1)} < \delta < 1$, where $\delta = a_1 p/(a_2 q)$.

Terminal state	Extremal control	Domain of controllability
$C_1 \begin{cases} x_1(t_1) = 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)\,(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 > s^2 - (b_2 x_2^0)^2$
$C_2 \begin{cases} x_1(t_1) = 0 \\ x_2(T) = 0 \\ y(T) > 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 > s^2 + (R-1)\,(b_2 x_2^0)^2$
$C_4 \begin{cases} x_1(t_2) > 0 \\ x_2(T) = 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_2 \\ 0 \text{ for } t_2 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 \geq R\{s^2 - (b_1 x_1^0)^2\}$ $a_1 b_1 y_0^2 \leq s^2 + A(b_2 x_2^0)^2$
$C_5 \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = 0 \text{ for } 0 \leq t \leq T$	$a_1 b_1 y_0^2 \leq Rs^2\{1 - 1/z^2\}$ $a_1 b_1 y_0^2 < R\{s^2 - (b_1 x_1^0)^2\}$
$C_5^S \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq T - \tau_1 \\ 0 \text{ for } T - \tau_1 < t \geq T \end{cases}$	$a_1 b_1 y_0^2 > Rs^2\{1 - 1/z^2\}$ $a_1 b_1 y_0^2 > s^2 + A(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 < s^2 + B(b_2 x_2^0)^2$

Definition of times:

(a) t_1 is first t such that $x_1(t_1) = 0$.

(b) t_2 is first t such that $2b_1 x_1(t_2) x_2^0 + b_2 (x_2^0)^2 + a_2 y^2(t_2)$.

(c) τ_1 is determined by $\cosh \sqrt{a_2 b_2}\, \tau_1 = \dfrac{R - \delta}{R - 1}$.

analysis (step (e) of the solution procedure outlined in [28]) yields Table II, which was previously obtained [28] by unsupported means. Details of this further analysis are given in Section 7.

It seems appropriate at this time to point out to the reader that the optimal control has been expressed as an open-loop control. For an open-loop control $u = u(t;\, x_0,\, t_0)$ one specifies the control as a function of time t during the length of the planning horizon $0 \leq t \leq T$. Such a control only depends on the parameter t *and* the initial conditions x_0, t_0. This control is not directly influenced by the current state of the system. On the other hand, one can consider a closed-loop (or feedback) control $u = k(x,\, t)$ which depends upon the current state of the system. The results presented in [14] are expressed in the form of a closed-loop control. For one-sided problems with deterministic systems (i.e., dynamics), it is well known that open-loop control and closed-loop control yield identical results in trajectory and payoff [11]. It is, of course, a simple matter to convert the optimal open-loop control presented in Table II into a closed-loop control.

It also seems appropriate to note the following for the quantities A and B which appear in inequalities defining various domains of controllability:

(2) (a) for $0 \leq \delta < R - \sqrt{R(R-1)}$, we have $A > B > 0$,

(3) (b) for $\delta = R - \sqrt{R(R-1)}$, we have $A = B = 0$,

TABLE II. *Solution to Isbell-Marlow Problem for* $R - \sqrt{R(R-1)} < \delta < 1$

Nonrestrictive Assumption: $R > 1$, i.e., $a_1 b_1 > a_2 b_2$.

Case (2): $R - \sqrt{R(R-1)} < \delta < 1$, where $\delta = a_1 p/(a_2 q)$.

Terminal state	Optimal control	Region of initial force levels
$C_1 \begin{cases} x_1(t_1) = 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)\ (b_2 x_2^0)^2$ $a_1 b_1 y_0^2 \geq s^2 + B(b_2 x_2^0)^2$
$C_2 \begin{cases} x_1(t_1) = 0 \\ x_2(T) = 0 \\ y(T) > 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 > s^2 + (R-1)\ (b_2 x_2^0)^2$
$C_4 \begin{cases} x_1(t_2) > 0 \\ x_2(T) = 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_2 \\ 0 \text{ for } t_2 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 \geq R\{s^2 - (b_1 x_1^0)^2\}$ $a_1 b_1 y_0 \leq s^2 + A(b_2 x_2^0)^2$
$C_5 \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = 0$ for $0 \leq t \leq T$	$a_1 b_1 y^2 \leq Rs^2\{1 - 1/z^2\}$ $a_1 b_1 y_0 < R\{s^2 - (b_1 x_1^0)^2\}$
$C_5^S \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq T - \tau_1 \\ 0 \text{ for } T - \tau_1 < t \geq T \end{cases}$	$a_1 b_1 y_0^2 > Rs^2\{1 - 1/z^2\}$ $a_1 b_1 y_0^2 > s^2 + A(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 < s^2 + B(b_2 x_2^0)^2$

Definitions of times: for t_1, t_2, and τ_1, see Table I.

(4) (c) for $R - \sqrt{R(R-1)} < \delta < 1$, we have $A < B < 0$.

Furthermore, let us recall that A and B are defined by (see also Section 3).

(5) $$A = A(R, z) = \frac{R(z^2 - 1) - z^2}{(z-1)^2},$$

(6) $$B = B(R, z) = \frac{R(z^2 - 1) - z^2}{z^2}.$$

The quantity z, defined by $z = (R - \delta)/(R - 1)$, relates A and B to δ.

For Case (3): $\delta = R - \sqrt{R(R-1)}$, the solution is shown in Table III. These results extend those given in [28]. The reader should note the reduction in dimension (from three to two) for the domain of controllability for C_5^S, $D(C_5^S)$. Case (3) is an important case for understanding the solution to the Isbell-Marlow problem. For $R - \sqrt{R(R-1)} < \delta < 1$, the domains of controllability $D(C_4)$, $D(C_5)$, and $D(C_5^S)$ do not overlap. However, for $0 \leq \delta < R - \sqrt{R(R-1)}$, this is no longer true. When $\delta = R - \sqrt{R(R-1)}$, any point $P^0 = (x_1^0 \quad x_2^0 \quad y_0)$ such that $a_1 b_1 y_0^2 = s^2$ belongs to each of $D(C_4)$, $D(C_5)$, and $D(C_5^S)$. Furthermore, these are the only such points. For $R - \sqrt{R(R-1)} < \delta < 1$ both optimal trajectories and the optimal control are unique. In Case (3): $\delta = R - \sqrt{R(R-1)}$, optimal trajectories are unique for all terminal states except when $a_1 b_1 y_0^2 = s^2$. Furthermore, extremals leading

TABLE III. *Solution to Isbell-Marlow Problem for* $\delta = R - \sqrt{R(R-1)}$

Nonrestrictive assumption: $R > 1$, i.e., $a_1 b_1 > a_2 b_2$.

Case (3): $\delta = R - \sqrt{R(R-1)}$, where $\delta = a_1 p/(a_2 q)$.

Terminal state	Optimal control	Region of initial force levels
$C_1 \begin{cases} x_1(t_1) = 0 \\ x_2(T) > 0 \\ v(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 \geq s^2$
$C_2 \begin{cases} x_1(t_1) = 0 \\ x_2(T) = 0 \\ y(T) > 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 > s^2 + (R-1)(b_2 x_2^0)^2$
$C_4 \begin{cases} x_1(t_2) > 0 \\ x_2(T) = 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_2 \\ 0 \text{ for } t_2 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 \geq R\{s^2 - (b_1 x_1^0)^2\}$ $a_1 b_1 y_0^2 \leq s^2$
$C_5 \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = 0$ for $0 \leq t \leq T$	$a_1 b_1 y_0^2 < R\{s^2 - (b_1 x_1^0)^2\}$ $a_1 b_1 y_0^2 \leq s^2$
$C_5^S \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_3 < t_1 \\ 0 \text{ for } t_3 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 = s^2$ $b_2 x_2^0 > b_1 x_1^0 (\sqrt{R/(R-1)} - 1)$

Definition of times:

(a) t_3 is *any* time such that $0 \leq t_3 < t_1$.

(b) For t_1, t_2, and τ_1, see Table I.

to C_5^S are not unique. Moreover, for C_5^S, the optimal control is no longer unique: for any $P^0 \epsilon D(C_5^S)$ any policy as shown in Table III with $t_3 < t_1$ leads to the same payoff, $P_5^S = -qx_2^0\sqrt{(R-1)/R}$. Furthermore, any initial point $P^0 \epsilon D(C_1)$ or $D(C_4)$ or $D(C_5)$ with $a_1 b_1 y_0^2 = s^2$ also leads to exactly the same payoff when an optimal control is used, i.e., for such P^0 we have $P_1 = P_4 = P_5 = P_5^S$.

For Case (4): $0 \leq \delta < R - \sqrt{R(R-1)}$, the domains of controllability and corresponding extremal controls are shown in Table IV. As previously for $R - \sqrt{R(R-1)} \leq \delta < 1$, $D(C_1)$ overlaps $D(C_4)$, $D(C_5)$, and $D(C_5^S)$. However, for $0 \leq \delta < R - \sqrt{R(R-1)}$, $D(C_4)$, $D(C_5)$, and $D(C_5^S)$ overlap each other (i.e., they are not disjoint sets) (as careful study of this table will show). (Inequalities, such as $A(R, z) > B(R, z)$ for $0 \leq \delta < R - \sqrt{R(R-1)}$ (which are essential in such considerations) are given in [28].)

A major difference between the results presented here (Table IV) and those of [28] is that extremals do lead to C_5^S. However, this gap in our earlier analysis (see [28]) has had no effect upon the solution which we presented earlier, since it may be shown for an initial point $P^0 \epsilon D(C_5) \cap D(C_5^S)$ we have $P_5(P^0) \geq P_5^S(P^0)$ when the corresponding extremal controls are used. Hence, extremals leading to C_5^S may be dropped (as they were previously inadvertently omitted) from further consideration. Again additional considerations in the large (step (e) of the general solution procedure presented in [28]) must be used to determine the optimal policy, and this yields the solution shown in Table V. However the details are considerably more complicated this time. (For $R - \sqrt{R(R-1)} < \delta < 1$, the payoff

TABLE IV. *Extremals for Isbell-Marlow Problem for* $0 \leq \delta < R - \sqrt{R(R-1)}$

Nonrestrictive assumption: $R > 1$, i.e., $a_1 b_1 > a_2 b_2$.
Case (4): $0 \leq \delta < R - \sqrt{R(R-1)}$, where $\delta = a_1 p/(a_2 q)$.

Terminal state	Extremal control	Domain of controllability
$C_1 \begin{cases} x_1(t_1) = 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)\,(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 > s^2 - (b_2 x_2^0)^2$
$C_2 \begin{cases} x_1(t_1) = 0 \\ x_2(T) = 0 \\ y(T) > 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_1 \\ 0 \text{ for } t_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 > s^2 + (R-1)\,(b_2 x_2^0)^2$
$C_4 \begin{cases} x_1(t_2) > 0 \\ x_2(T) = 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq t_2 \\ 0 \text{ for } t_2 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)\,(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 \geq R\{s^2 - (b_1 x_1^0)^2\}$ $a_1 b_1 y_0^2 \leq s^2 + A(b_2 x_2^0)^2$
$C_5 \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = 0 \text{ for } 0 \leq t \leq T$	$a_1 b_1 y_0^2 < s^2 + (R-1)\,(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 \leq Rs^2\{1 - 1/z^2\}$ $a_1 b_1 y_0^2 < R\{s^2 - (b_1 x_1^0)^2\}$
$C_5^S \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$\phi^*(t) = \begin{cases} 1 \text{ for } 0 \leq t \leq T - \tau_1 \\ 0 \text{ for } T - \tau_1 < t \leq T \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)\,(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 < Rs^2\{1 - 1/z^2\}$ $a_1 b_1 y_0^2 > s^2 + A(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 > s^2 + B(b_2 x_2^0)^2$

Definition of times: for t_1, t_2, and τ_1, see Table I.

TABLE V. *Solution to Isbell-Marlow Problem for* $0 \leq \delta < R - \sqrt{R(R-1)}$

Nonrestrictive assumption: $R > 1$, i.e., $a_1 b_1 > a_2 b_2$.

Case (4): $0 \leq \delta < R - \sqrt{R(R-1)}$, where $\delta = a_1 p/(a_2 q)$.

Terminal state	Region of initial force levels
$C_1 \begin{cases} x_1(t_1) = 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 \geq Rs^2 - R\{b_1 x_1^0[z^2(R-1) + R]/(2R) + b_2 x_2^0\}^2/z^2$ for $0 \leq x_1^0 < (b_2 x_2^0)/(kb_1)$ $a_1 b_1 y_0^2 \geq h(P^0, R, z)$ for $0 \leq x_2^0 \leq kb_1 x_1^0/b_2$
$C_2 \begin{cases} x_1(t_1) = 0 \\ x_2(T) = 0 \\ y(T) > 0 \end{cases}$	$a_1 b_1 y_0^2 > s^2 + (R-1)(b_2 x_2^0)^2$
$C_4 \begin{cases} x_1(t_2) > 0 \\ x_2(T) = 0 \\ y(T) = 0 \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 \geq R\{s^2 - (b_1 x_1^0)^2\}$ $a_1 b_1 y_0^2 \leq h(P^0, R, z)$ for $0 \leq x_2^0 \leq kb_1 x_1^0/b_2$
$C_5 \begin{cases} x_1(T) > 0 \\ x_2(T) > 0 \\ y(T) = 0 \end{cases}$	$a_1 b_1 y_0^2 < s^2 + (R-1)(b_2 x_2^0)^2$ $a_1 b_1 y_0^2 < R\{s^2 - (b_1 x_1^0)^2\}$ $a_1 b_1 y_0^2 \leq Rs^2 - R\{b_1 x_1^0[z^2(R-1) + R]/(2R) + b_2 x_2^0\}^2/z^2$ for $0 \leq x_1^0 < (b_2 x_2^0)/(kb_1)$

Explanation of symbols:

(a) For t_1, t_2, and τ_1, see Table I.

(b) $k = [z^2 - R(z-1)^2]/(2R)$.

associated with an extremal leading to C_1 was always strictly dominated by the payoff associated with any other extremal when there were multiple extremals leading to the terminal surface from an initial point P^0.)

For Case (4): $0 \leq \delta < R - \sqrt{R(R-1)}$, the solution is shown in Table V. It is of interest to note that the solutions contain a dispersal surface (see pp. 134–141 in [13]), a solution aspect rarely encountered in optimal control problems (see [7] for the only example of which we are aware, other than those given in Isaacs' book [13]). The reader should note that the solution shown in Table V is identical with that given earlier (see Table II of [28]).

A considerable amount of effort has gone into the study of extremals for $0 \leq \delta < R - \sqrt{R(R-1)}$. In Table VI we show event (such as annihilation of X_1, end of battle, etc.) times and payoffs for Case (4): $0 \leq \delta < R - \sqrt{R(R-1)}$. We omit derivation of these times, which is readily done using elementary considerations (such as combination of the extremal control with the solution to the force level equations). The solution to the force level equations is consequently shown in Table VII. Additionally, the reader should note that if we have $\phi(t) =$ constant for all $t \epsilon [t_1, t_2]$, then the following "generalized square law" holds

$$\zeta^2(t=t_1) - \zeta^2(t=t_2) = \{\phi a_1 b_1 + (1-\phi) a_2 b_2\} \{y^2(t=t_1) - y^2(t=t_2)\}, \tag{7}$$

where

$$\zeta(t) = b_1 x_1(t) + b_2 x_2(t).$$

This generalized square law has been used (along with the extremal control to C_i) to express $P_i = ry(T) - px_1(T) - qx_2(T)$ in the form shown in Table VI.

As a final check on our theoretical developments, we had some numerical computations done for Case (4). The results shown in Table VI have been used in this work. We would like to thank a M.S. thesis student Robert L. Powers, LCDR, USN, for his efforts in this area. All Powers' computations have supported the solution shown in Table V. Finally, because in our earlier work [28] we had erroneously concluded that no extremals led to C_5^S for $0 \leq \delta < R - \sqrt{R(R-1)}$, we have verified that extremals do lead to C_5^S by examining the time history of the force levels computed according to the extremal control.

TABLE VI. *Event Times and Payoffs for Isbell-Marlow Problem when* $0 \leq \delta < R - \sqrt{R(R-1)}$ – Continued

Terminal state C_1: $x_1(t_1) = 0$, $x_2(T) > 0$, $y(T) = 0$

$$t_1 = \frac{1}{\sqrt{a_1 b_1}} \ln \left\{ \frac{\sqrt{a_1 b_1 y_0^2 - s^2 + (b_2 x_2^0)^2} - \beta_2 x_2^0}{\sqrt{a_1 b_1}\, y_0 - s} \right\}$$

$$T = t_1 + \frac{1}{\sqrt{a_2 b_2}} \tanh^{-1} \left\{ \frac{\sqrt{a_1 b_1 y_0^2 - s^2 + (b_2 x_2^0)^2}}{b_2 x_2^0 \sqrt{R}} \right\}$$

$$P_1 = \frac{-q}{b_2 \sqrt{R}} \sqrt{s^2 + (R-1)(b_2 x_2^0)^2 - a_1 b_1 y_0^2}$$

Terminal state C_2: $x_1(t_1) = 0$, $x_2(T) = 0$, $y(T) > 0$

$$t_1 = \frac{1}{\sqrt{a_1 b_1}} \ln \left\{ \frac{\sqrt{a_1 b_1 y_0^2 - s^2 + (b_2 x_2^0)^2} - b_2 x_2^0}{\sqrt{a_1 b_1}\, y_0 - s} \right\}$$

$$T = t_1 + \frac{1}{\sqrt{a_2 b_2}} \tanh^{-1} \left\{ \frac{b_2 x_2^0 \sqrt{R}}{\sqrt{a_1 b_1 y_0^2 - s^2 + (b_2 x_2^0)^2}} \right\}$$

Table VI—Continued

Terminal state C_4: $x_1(t_2)=x_1(T)>0$, $x_2(T)=0$, $y(T)=0$

t_2:

for $a_1b_1y_0^2>s^2$

$$t_2=\frac{1}{\sqrt{a_1b_1}}\ln\left\{\frac{y(t_2)-\sqrt{y^2(t_2)-y_0^2+s^2/(a_1b_1)}}{y_0-s/\sqrt{a_1b_1}}\right\}$$

for $a_1b_1y_0^2<s^2$

$$t_2=\frac{1}{\sqrt{a_1b_1}}\ln\left\{\frac{\sqrt{y^2(t_2)-y_0^2+s^2/(a_1b_1)}-y(t_2)}{s/\sqrt{a_1b_1}-y_0}\right\}$$

for $a_1b_1y_0^2=s$

$$t_2=\frac{1}{\sqrt{a_1b_1}}\ln\left\{\frac{y_0}{y(t_2)}\right\}$$

where

$$y(t_2)=\sqrt{\frac{x_2^0}{a_2}\{(2R-1)b_2x_2^0+2\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}\}},$$

also

$$T=t_2+\frac{1}{\sqrt{a_2b_2}}\cosh^{-1}\left\{\frac{Rb_2x_2^0+\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}}{(R-1)b_2x_2^0+\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}}\right\}$$

$$P_4=\frac{-q\delta}{b_2R}\{b_2x_2^0(R-1)+\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}\}$$

Terminal state C_5: $x_1(T)>0$, $x_2(T)>0$, $y(T)=0$

(no switch: $\phi^*(t)=0$ *for* $0\le t\le T$)

$$T=\frac{1}{\sqrt{a_2b_2}}\tanh^{-1}\left\{\frac{\sqrt{a_2b_2}y_0}{b_1x_1^0+b_2x_2^0}\right\}$$

Terminal state C_5^S: $x_1(T)>0$, $x_2(T)>0$, $y(T)=0$

$$\left(\text{switch in tactics: } \phi^*(t)=\begin{cases}1 \text{ for } 0\le t\le T-\tau_1\\ 0 \text{ for } T-\tau_1<t\le T\end{cases}\right)$$

$$\tau_1=\frac{1}{\sqrt{a_2b_2}}\cosh^{-1}\left(\frac{R-\delta}{R-1}\right)$$

$$T=\tau_1+\frac{1}{2\sqrt{a_1b_1}}\ln\left\{\frac{(\sqrt{R(z^2-1)}-z)(\sqrt{a_1b_1}y_0+s)}{(\sqrt{a_1b_1}y_0-s)(\sqrt{R(z^2-1)}+z)}\right\}$$

$$P_5^S=\frac{q}{b_2}\left\{-z(b_2x_2^0)\left(\frac{R-1}{R}\right)+\frac{1}{R}\sqrt{(z^2(R-1)-R)(a_1b_1y_0^2-s^2)}\right\}$$

TABLE VII. *Solution to Force Level Equations for Two-on-One Combat*

Combat Equations:

$$\frac{dx_1}{dt} = -\phi a_1 y \qquad \text{with } x_1(t=0) = x_1^0,$$

$$\frac{dx_2}{dt} = -(1-\phi)a_2 y \qquad \text{with } x_2(t=0) = x_2^0,$$

$$\frac{dy}{dt} = -b_1 x_1 - b_2 x_2 \qquad \text{with } y(t=0) = y_0.$$

Solution: when $\phi(t) = \text{constant}$ for $t \in [0, t_1]$, then

$$x_1(t) = x_1^0 + \phi a_1 \left\{ \frac{(b_1 x_1^0 + b_2 x_2^0)}{[\phi a_1 b_1 + (1-\phi)a_2 b_2]} \cosh \sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}\, t \right.$$
$$\left. - \frac{y_0}{\sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}} \sinh \sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}\, t - \frac{(b_1 x_1^0 + b_2 x_2^0)}{[\phi a_1 b_1 + (1-\phi)a_2 b_2]} \right\}$$

$$x_2(t) = x_2^0 + (1-\phi)a_2 \left\{ \frac{(b_1 x_1^0 + b_2 x_2^0)}{[\phi a_1 b_1 + (1-\phi)a_2 b_2]} \cosh \sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}\, t \right.$$
$$\left. - \frac{y_0}{\sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}} \sinh \sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}\, t - \frac{(b_1 x_1^0 + b_2 x_2^0)}{[\phi a_1 b_1 + (1-\phi)a_2 b_2]} \right\}$$

$$y(t) = y_0 \cosh \sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}\, t - \frac{(b_1 x_1^0 + b_2 x_1^0)}{\sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}} \sinh \sqrt{\phi a_1 b_1 + (1-\phi)a_2 b_2}\, t$$

5. NECESSARY CONDITIONS OF OPTIMALITY

We will use Speyer's approach* [27] (see also [15]) of adjoining the state-variable constrain directly to the criterion functional with a Lagrange multiplier. (In Appendix B we relate this method t the one previously used in [31].) In Speyer's approach [27], the Hamiltonian is given by

$$H(t, x_i, p_i, \phi) = -p_1 \phi a_1 y - p_2 (1-\phi) a_2 y - p_3 (b_1 x_1 + b_2 x_2) + \eta_1(t) x_1 + \eta_2(t) x_2, \tag{8}$$

where

$$\eta_i(t) \begin{cases} = 0 & \text{for } x_i > 0, \\ \geqslant 0 & \text{for } x_i = 0. \end{cases}$$

The adjoint system of differential equations for the dual variables is

$$\frac{dp_1}{dt} = -\frac{\partial H}{\partial x_1}(t, x_i, p_i, \phi^*) = b_1 p_3 - \eta_1(t), \tag{9}$$

*This is apparently only the second application of these important results to a variational problem in operations research. Th only other application known to the author is [35].

$$\frac{dp_2}{dt} = -\frac{\partial H}{\partial x_2}(t, x_i, p_i, \phi^*) = b_2 p_3 - \eta_2(t), \tag{10}$$

$$\frac{dp_3}{dt} = -\frac{\partial H}{\partial y}(t, x_i, p_i, \phi^*) = \phi^* a_1 p_1 + (1-\phi^*) a_2 p_2. \tag{11}$$

Deferring the discussion of boundary conditions for the dual variables until later, we note the transversality condition

$$H(t=T, x_i, p_i, \phi^*) = 0. \tag{12}$$

When $x_1, x_2 > 0$, the maximum principle yields the extremal control [28]

$$\phi^*(t) = \begin{cases} 1 & \text{for } v(t) > 0, \\ \\ 0 & \text{for } v(t) < 0, \end{cases} \tag{13}$$

where $v(t) = (-p_1)a_1 - (-p_2)a_2$. In [28] we showed that there are no singular subarcs (see ch. 8 in [4]) in the solution.

Without loss of generality, let us consider a constrained subarc on which $x_1(t) = 0$ for $t_1 \leq t \leq T$ (and x_2, $y > 0$ for $t < T$). Since we have $dx_1/dt = 0$, the control is clearly $\phi^*(t) = 0$ for $t_1 \leq t \leq T$. The condition that $\partial H/\partial\phi = 0$ yields that the following relationship holds between the dual variables on the constrained subarc*

$$a_1 p_1(t) = a_2 p_2(t), \tag{14}$$

since

$$\frac{\partial H}{\partial \phi} = y(a_1 p_1 - a_2 p_2).$$

To determine the multiplier $\eta_1(t)$, we compute

$$\frac{d}{dt}\left(\frac{\partial H}{\partial \phi}\right) = 0 = y\left(a_1 \frac{dp_1}{dt} - a_2 \frac{dp_2}{dt}\right) + \frac{dy}{dt}(a_1 p_1 - a_2 p_2)$$

and by using (9), (10), and (14), this yields

---

*Strictly speaking, this result depends upon the condition that $y(t) > 0$. We note that Speyer's results (see Problem 1 in Appendix B) have been derived under the assumption that $(C)_u \neq 0$ along an optimal trajectory [15]. For the problem at hand this requirement translates to $(\dot{x}_1)_\phi = -a_1 y \neq 0$. Thus, when Y loses, (14) is not a necessary condition for $\partial H/\partial\phi = 0$ at $t = T$, when $y(T) = 0$. However, since $y(t) > 0$ for $t < T$, we see that the condition $(\dot{x}_1)_\phi \neq 0$ is violated in this case at only an isolated point in time. This is equivalent to the control being undetermined at the isolated point, the end of battle when Y (who controls ϕ) is annihilated. It is clear that this does not affect the state trajectory (namely, the SVIC is not violated). Moreover, the dual variables (p_i's) in Speyer's approach are continuous everywhere [33] (see also [15]). Thus, by invoking the continuity of the dual variables (14) may be justified when $y(T) = 0$.

$$\eta_1(t) = \frac{p_3(t)}{a_1}(a_1b_1 - a_2b_2). \tag{15}$$

The interpretation of $\eta_1(t)$ (see Appendix B for further discussion) is the rate of marginal return to Y for keeping $x_1 = 0$. Thus, (intuitively) Y tries to annihilate X_1 only when it profits him to do so. Furthermore, the requirement that $\eta_1(t) \geq 0$ when $x_1 = 0$ for a finite interval of time yields that we must nave

$$a_1b_1 \geq a_2b_2, \tag{16}$$

since it may be shown that $p_3(t) > 0$ for $t < T$. The nonrestrictive assumption that $a_1b_1 > a_2b_2$ (i.e., $R > 1$) implies that it is nonoptimal to have $x_2 = 0$ for a finite interval of time. Hence, we may drop C_3 (see section 2.) from further consideration.

Furthermore, there are corner conditions* that must be satisfied at entry to a constrained subarc. (There are also exit conditions. However, for the model at hand once such a constrained subarc has been entered, it cannot be left, since without replacements force levels cannot increase. In other words, the state boundary is "absorbing" (as we first coined the word in [33]).) Let t_1 denote the time of entry to a constrained subarc, and let t_1^- denote a left-hand limit. Then we have [19], [33]

$$p_i(t_1^-) = p_i(t_1^+) \qquad \text{for } i = 1, 2, 3. \tag{17}$$

Thus, when the necessary conditions of optimality are expressed in the form considered by Speyer and Bryson [27] (see Appendix B for a further discussion) the adjoint variables are continuous across all corners† (both interior to and on the boundary of the state space). This is one advantage of using Speyer's approach.

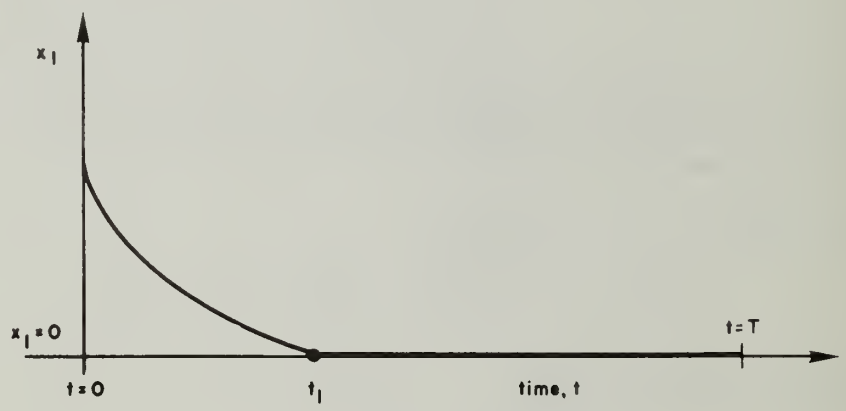

FIGURE 1. Entry to constrained subarc

Additionally, we have

$$H(t_1^-) = H(t_1^+), \tag{18}$$

*For a discussion of corner conditions for corner conditions for corners at interior points of the state space, see [31].

† This is only true, however, for a first order SVIC.

where $H(t_1^-)$ denotes $H(t_1, x_i(t_1), p_i(t_1^-), \phi^*(t_1^-))$, and this yields

$$a_1 p_1(t_1^-) = a_1 p_1(t_1) = a_2 p_2(t_1^-) = a_2 p_2(t_1). \tag{19}$$

Let us finally consider the boundary conditions for the dual variables at $t=T$. As seen above, the nonrestrictive assumption that $a_1 b_1 >_1 a_2 b_2$ yields that no extremals lead to C_3. For C_2: $x_1(T)=x_2(T)=0$, $y(T)>0$, we have [4]

$$p_1(T) = -p + \nu_1, \qquad p_2(T) = -q + \nu_2, \qquad p_3(T) = r, \tag{20}$$

since T is determined by the (equality) terminal constraints $x_1(T)=0$ and $x_2(T)=0$. Since these are equality constraints, the multipliers ν_1 and ν_2 are unrestricted in sign. Since T is unspecified, the transversality condition (12) with $\phi^*(T)=0$ yields that $-p_2(T)a_2 y=0$ so that $p_2(T)=0$ and $\nu_2=q$. The condition (14) which, in particular, holds at $t=T$ yields $p_1(T)=0$. Thus, we have for $C_2[x_1(T)=0$ before: $x_2(T)=0$, $y(T)>0]$:

$$p_1(T) = 0, \qquad p_2(T) = 0, \qquad p_3(T) = r. \tag{21}$$

The three terminal states C_1, C_4, and C_5 may be discussed collectively. In all three cases, T is determined by the terminal equality constraint $y(T)=0$ so that the transversality condition (12) yields $p_3(T)=0$. Then, according to the results presented in [33], we have for C_1, C_4, and C_5:

$$p_1(T) = -p + \nu_1, \qquad p_2(T) = -q + \nu_2, \qquad p_3(T) = 0, \tag{22}$$

where

$$\nu_i \begin{cases} = 0 & \text{for } x_i(T) > 0, \\ \geqslant 0 & \text{for } x_i(T) = 0 \text{ but } x_i(t) > 0 \text{ for } t < T, \\ \text{unrestricted} & \text{for } x_i(T) = 0 \text{ and } x_i(t) = 0 \\ & \text{for } t_i \leqslant t \leqslant T \text{ with } t_i < T. \end{cases}$$

The latter condition that, for example, the multiplier ν_1 is unrestricted when the system is on a constrained subarc for a finite interval of time is because the boundary of the state space is "absorbing" so that the state constraint $x_1 \geqslant 0$ essentially acts like a terminal equality constraint as far as the determination of boundary conditions for the adjoint variables [33]. If there were replacements in the model (1) so that the boundary of the state space would not be "absorbing," then we would have $\nu_i \geqslant 0$ for $x_i(T)=0$.

6. SYNTHESIS OF EXTREMAL CONTROL

In this section we sketch how the above necessary conditions of optimality are used to construct extremals. In all cases, for C_2 we have that $\phi^*(t)=0$ for $t_1 \leqslant t \leqslant T$. Furthermore, by (19) we have

$$v(t_1) = (-p_1(t_1))a_1 - (-p_2(t_1))a_2 = 0. \tag{23}$$

Now using (9) and (10), we readily find that

$$dv/dt = p_3(t)\,(-a_1b_1 + a_2b_2) < 0, \tag{24}$$

since $p_3(t) > 0$. Hence, $v(t) > 0$ for $t < t_1$, and we have $\phi^*(t) = 1$ for all $t \in [0, t_1)$ by (13).

For C_1, C_4, and C_5, we must distinguish two cases: (1) $a_1p \geq a_2q$, and (2) $a_1p < a_2q$. We discuss each case separately.

6.1. Extremal Control for $a_1p \geq a_2q$

For $C_1[x_1(T) = 0,\ x_2(T) > 0,\ y(T) = 0]$: when the system has been on a constrained subarc for a finite interval of time, we have by (14) and (22)

$$a_1(-p + \nu_1) = -a_2q, \text{ or}$$

$$\nu_1 = (a_1p - a_2q)/a_1. \tag{25}$$

However, since ν_1 is unrestricted due to the "absorbing" boundary of the state space, for the problem at hand no useful information is to be obtained from (25). It should be noted that (23) and (24) again hold so that $\phi^*(t) = 1$ for $0 \leq t < t_1$. When $x_1(T) = 0$, but $x_1(t) > 0$ for $t < T$, we have $\phi^*(t) = 1$ for $T - \delta_1 \leq t < T$, where $\delta_1 > 0$ so that by (13) and (24), we must have $v(t = T) \geq 0$. Combining this with (22), we obtain

$$a_1(p - \nu_1) - a_2q \geq 0, \text{ or}$$

$$\nu_1 \leq (a_1p - a_2q)/a_1. \tag{26}$$

By (22), we must have $\nu_1 \geq 0$ so that it is consistent with an optimal policy to have $x_1(T) = 0$, but $x_1(t) > 0$ for $t < T$. In either case, we have

$$\phi^*(t) = \begin{cases} 1 & \text{for } 0 \leq t \leq t_1, \\ 0 & \text{for } t_1 < t \leq T, \text{ with } t_1 \leq T. \end{cases} \tag{27}$$

For $C_4[x_1(T) > 0,\ x_2(T) = 0,\ y(T) = 0]$: by our assumption that $a_1b_1 > a_2b_2$ we saw in Section 5 that it is nonoptimal to have $x_2 = 0$ for a finite interval of time. Hence, for extremals to reach C_4 we must have $x_2(t) > 0$ for $t < T$. Clearly, we must then have $\phi^*(t) = 0$ for $T - \delta_1 \leq t \leq T$, where $\delta_1 > 0$ so that by (13) and (24) we must have $v(t = T) < 0$. Combining this with (22), we obtain

$$a_1p - a_2(q - \nu_2) < 0, \text{ or}$$

$$\nu_2 < (a_2q - a_1p)/a_2. \tag{28}$$

However, this implies $\nu_2 < 0$ (since $a_2q - a_1p \leq 0$) so that by (22) it is nonoptimal to reach C_4 in this case.

The arguments for C_5 are the same as given in [28] and omitted.

6.2. Extremal Control for $a_1p < a_2q$

For C_1, (26) yields $\nu_1 < 0$ so that by (22) it is nonoptimal to have $x_1(t) = 0$, but $x_1(t) > 0$ for $t < T$. Thus, (27) can only hold with $t_1 < T$.

Extremals can now reach C_4 with $x_2(t) > 0$ for $t < T$, since in this case (28) can be satisfied with $\nu_2 \geq 0$. Moreover, from $v(T) < 0$ and (22), we have

$$\frac{a_1p}{a_2} < (-p_2(t=T)) \leq q. \tag{29}$$

The latter of these two inequalities was conjectured in [28]. Thus, our recent work using the theory of SVIC's has proved this conjecture to be true. The multiplier ν_2 is chosen so that $x_2(T)=0$. By using a result given in [28] for $p_2(T)$, it is readily shown that

$$\nu_2 = q\left\{1-\delta-\frac{b_2x_2^0\delta(R-1)}{\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}}\right\}. \tag{30}$$

7. DETERMINATION OF THE OPTIMAL CONTROL

It remains to discuss how the optimal fire distribution policy is determined from the extremal controls whose development was outlined in the previous section. This comprises steps (c) through (e) of the general solution procedure given in [28]. The basic idea may be summarized as follows: given than an optimal control (or policy) exists, then (1) if extremals are unique,* the extremal control is optimal; or (2) if there are multiple extremals (but a finite number), the optimal control is obtained by determining the largest value of the return functional corresponding to these extremals. It cannot be too strongly emphasized that the approach vitally depends upon the existence of an optimal control.† Moreover, in [28] it was implicitly assumed that an optimal policy exists for the problem (1). Unfortunately, this is not true, in general, for variational problems arising in operations research (see [6]). Thus, to validate out results for the Isbell-Marlow fire programming problem we must show the existence of an optimal control. This readily follows, however, from the control variable appearing linearly in (1) and the uniform boundedness** of responses to the controller ϕ by Corollary 2 on p. 262 of [17].

As shown in [28], the definitions of the terminal states (primarily the nonnegativity of the force levels) may be combined with the corresponding extremal controls to determine the domains of controllability. For example, the extremal control (27) is combined with the requirement †† that $y(t=t_1) > 0$ (as must hold for $\delta < 1$) to yield

$$a_1b_1y_0^2 > s^2-(b_2x_2^0)^2, \tag{31}$$

which partially defines $D(C_1)$ and appears in Table I. The development of results that were given in

*By this we mean that from any initial point P^0 there is only one extremal leading to the terminal surface.

†In [6] we give an example of the above approach failing when an optimal policy doesn't exist.

**This is a consequence of the facts that force levels (state variables) must be nonnegative and force levels cannot increase over time.

††The time t_1 is defined as the smallest time such that $x(t=t_1)=0$.

[28] is omitted. However, the following new results (which extend those given in [28]) are developed in subsequent sections of the paper at hand:

(1) extremals to C_5^S for $0 \leq \delta < R - \sqrt{R(R-1)}$,
(2) the dominated payoff from extremals to C_5^S for $0 \leq \delta < R - \sqrt{R(R-1)}$,
(3) development of conditions for extremals to reach C_4.

Besides their own intrinsic importance, these results provide insight into solution phenomena and methods that play a vital role in determining the optimal control of deterministic Lanchester attrition processes for more complex battle dynamics.

Finally, we discuss the determination of the optimal fire distribution policy for $R - \sqrt{R(R-1)} < \delta < 1$. Our previous discussion [28] of this point was inadequate (as was that of [14]). By Theorems A1, A2, and A3 (see Appendix A for further details) we have that for $R - \sqrt{R(R-1)} < \delta < 1$:

a. $P_4(P^0) > P_1(P^0)$ for all $P^0 \in \{D(C_1) \cap D(C_4)\}$,
b. $P_5(P^0) > P_1(P^0)$ for all $P^0 \in \{D(C_1) \cap D(C_5)\}$,
c. $P_5^S(P^0) > P_1(P^0)$ for all $P^0 \in \{D(C_1) \cap D(C_5^S)\}$.

Considering the above and the results given in Table I, we see that optimal paths lead to C_1 for

$$s^2 + B(b_2 x_2^0)^2 \leq a_1 b_1 y_0^2 < s^2 + (R-1)\,(b_2 x_2^0)^2. \tag{32}$$

Thus, the main point to be learned from our re-examination of this problem is that considerations "in the large" play a major role in determining its solution for $\delta < 1$. It is conjectured that this is because the boundary of the state space is "absorbing" (i.e., no replacements are allowed). The author further conjectures that this difficulty will be present in all such optimal control problems in the Lanchester theory of combat.

8. EXTREMALS LEADING TO C_5^S

In our earlier work [28] we erroneously concluded that for $0 \leq \delta < R - \sqrt{R(R-1)}$ there were no extremals leading to C_5^S. Let us now retrace our earlier development in a correct manner.

There are two subcases for entry to C_5. We consider the extremals for which

$$\phi^*(t) = \begin{cases} 1 & \text{for } 0 \leq t \leq T - \tau_1, \\ 0 & \text{for } T - \tau_1 < t \leq T. \end{cases} \tag{33}$$

For convenience, let us say that such an extremal leads to C_5^S. In [28] we showed that

$$b_2 x_2^0 = b_2 x_2(T) z + b_1 x_1(T - \tau_1)(z - 1), \tag{34}$$

and

(35) $$(b_2x_2(T) - b_2x_2^0)^2 = (a_1b_1y_0^2 - s^2)/A.$$

We require that the right-hand side of (35) be positive in order that C_5^S be reached. Thus, for $P^0 \in D(C_5^S)$ we must have

(36) $$a_1b_1y_0^2 < s^2 \qquad \text{for } A < 0,$$

(37) $$a_1b_1y_0^2 = s^2 \qquad \text{for } A = 0,$$

(38) $$a_1b_1y_0^2 > s^2 \qquad \text{for } A > 0,$$

where the reader should keep (2) through (4) in mind.

Now when $A < 0$ (corresponding to $R - \sqrt{R(R-1)} < \delta < 1$), (35) may be solved to yield

(39) $$b_2x_2(T) = b_2x_2^0 + (1-z)\sqrt{\frac{s^2 - a_1b_1y_0^2}{R - z^2(R-1)}}.$$

The requirement that $x_2(T) > 0$ then leads to

(40) $$a_1b_1y_0^2 > s^2 + A(b_2x_2^0)^2,$$

where we have made use of (5). Combination of (34) and (39) yields that

(41) $$b_1x_1(T-\tau_1) = -b_2x_2^0 + z\sqrt{\frac{s^2 - a_1b_1y_0^2}{R - z^2(R-1)}}.$$

The requirement that $x_1(T-\tau_1) > 0$ then leads to

(42) $$a_1b_1y_0^2 < s^2 + B(b_2x_2^0)^2,$$

where we have made use of (6). Furthermore, the time $T-\tau_1$ may be determined by solving the following equation (see Table VII) $y(T-\tau_1) = y_0 \cosh \sqrt{a_1b_1}\,(T-\tau_1) - \frac{s}{\sqrt{a_1b_1}} \sinh \sqrt{a_1b_1}\,(T-\tau_1)$ to yield

(43) $$T-\tau_1 = \frac{1}{\sqrt{a_1b_1}} \ln\left\{\frac{\sqrt{a_1b_1y^2(T-\tau_1) - a_1b_1y_0^2 + s^2} - \sqrt{a_1b_1}\,y(T-\tau_1)}{s - \sqrt{a_1b_1}\,y_0}\right\}.$$

The requirement that the quantity in brackets is greater than one leads to

(44) $$y_0 > y(T-\tau_1).$$

Now, the generalized square law (7) may be combined with the extremal control (33) and (34) to yield

$$\sqrt{a_1 b_1}\, y(T-\tau_1) = \sqrt{R(z^2-1)}\, \sqrt{\frac{s^2 - a_1 b_1 y_0^2}{R - z^2(R-1)}}. \tag{45}$$

(44) and (45) may be combined to yield

$$a_1 b_1 y_0^2 > R s^2 \{1 - 1/z^2\}. \tag{46}$$

Furthermore, (43) and (45) may be combined to yield

$$T - \tau_1 = \frac{1}{2\sqrt{a_1 b_1}} \ln \left\{ \frac{(z - \sqrt{R(z^2-1)})\,(s + \sqrt{a_1 b_1}\, y_0)}{(s - \sqrt{a_1 b_1}\, y_0)\,(z + \sqrt{R(z^2-1)})} \right\}. \tag{47}$$

Finally, (39) may be used to show that the payoff P_5^S corresponding to use of the extremal control (33) is given by

$$P_5^S = \frac{q}{b_2} \left\{ -z \left(\frac{R-1}{R} \right) b_2 x_2^0 + \frac{1}{R} \sqrt{(R - z^2(R-1))\,(s^2 - a_1 b_1 y_0^2)} \right\}. \tag{48}$$

When $A > 0$ (corresponding to $0 \leq \delta < R - \sqrt{R(R-1)}$), we must have $a_1 b_1 y_0^2 > s^2$ by (38) so that analysis similar to the above readily yields results similar to (40), (42), and (46), but with the inequalities reversed (see Table IV). In [28] we had erroneously concluded that this case was impossible.

Next, we consider the extremals for which

$$\phi^*(t) = 0 \qquad \text{for } 0 \leq t \leq T. \tag{49}$$

Let us adopt the terminology that such an extremal leads to C_5 with payoff P_5. In [28] we showed that

$$P_5 = \frac{q}{b_2} \left\{ z \left(\frac{R-1}{R} \right) b_1 x_1^0 - \sqrt{s^2 - a_1 b_1 y_0^2 / R} \right\}. \tag{50}$$

9. PROOF OF A DOMINANCE THEOREM

In this section we show that for $0 \leq \delta < R - \sqrt{R(R-1)}$ we need not consider extremals that lead to C_5^S, since the payoff associated with such an extremal path is exceeded (or dominated) by the payoff associated with an extremal leading to C_5. Hence, our previous oversight on extremals leading to C_5^S is of no consequence.

Our main result is stated below as Theorem 1. (For definitions of $D(C_5)$ and $D(C_5^S)$, see Table IV.

THEOREM 1: Assume that $0 \leq \delta < R - \sqrt{R(R-1)}$. (This is equivalent to $z > \sqrt{R/(R-1)}$.) Then $P^0 = (x_1^0,\ x_2^0,\ y_0) \in D(C_5) \cap D(C_5^S)$ implies that $P_5 \geq P_5^S$ with equality holding only when $a_1 b_1 y_0^2 = R s^2 (1 - 1/z^2)$.

PROOF: A. First we will show that

(51) $$s^2 < a_1 b_1 y_0^2 < R s^2 (1 - 1/z^2)$$

and

(52) $$\frac{R}{R-1} < z^2,$$

imply that $P_5^S < P_5$. For $P^0 \in D(C_5) \cap D(C_5^S)$ we have (postponing for now the case when equality holds) that (51) holds, which implies that

(53) $$R(z^2-1)s^2 - z^2 a_1 b_1 y_0^2 > 0.$$

Squaring (53), we may rearrange the result to obtain

(54) $$(2Rsz\sqrt{s^2 - a_1 b_1 y_0^2/R})^2 < \{R(z^2+1)s^2 - z^2 a_1 b_1 y_0^2\}^2.$$

By observing that $1 - 1/z^2 < 1 + 1/z^2$ and considering (51), it is readily seen that

(55) $$R(z^2+1)s^2 - z^2 a_1 b_1 y_0^2 > 0,$$

so that we may take square roots in (54) to obtain

(56) $$2Rsz\sqrt{s^2 - a_1 b_1 y_0^2/R} < R(z^2+1)s^2 - z^2 a_1 b_1 y_0^2.$$

Now it is easily seen that (54) may be rearranged to yield

(57) $$(z^2(R-1) - R)\ (a_1 b_1 y_0^2 - s^2) < \{z(R-1)s - R\sqrt{s^2 - a_1 b_1 y_0^2/R}\}^2.$$

The conditions $s^2 < a_1 b_1 y_0^2$ and $z^2 > R/(R-1)$ are readily combined to yield

(58) $$R^2(s^2 - a_1 b_1 y_0^2/R) < z^2(R-1)^2 s^2.$$

Recalling that for $P^0 \in D(C_5) \cap D(C_5^S)$ we have $a_1 b_1 y_0^2 < R s^2(1 - 1/z^2) < R s^2$, we may extract square roots in (58) to obtain

(59) $$z(R-1)s - R\sqrt{s^2 - a_1 b_1 y_0^2/R} > 0.$$

Considering (51), (52), and (59), we may extract square roots in (57) to obtain

(60) $$\frac{1}{R}\sqrt{(z^2(R-1) - R)\ (a_1 b_1 y_0^2 - s^2)} < z\left(\frac{R-1}{R}\right)(b_1 x_1^0 + b_2 x_2^0) - \sqrt{s^2 - a_1 b_1 y_0^2/R}.$$

Some additional manipulations on (60) now readily yield that (51) and (52) imply that $P_5 > P_5^S$.

B. By using (48) and (50), it is readily shown that $a_1 b_1 y_0^2 = R s^2(1 - 1/z^2)$ implies that $P_5 = P_5^S$. By A and B above, Theorem 1 is proved. Q.E.D.

10. DEVELOPMENT OF CONDITIONS FOR EXTREMALS TO REACH C_4

Another topic for which our previous discussion [28] was inadequate was the development of conditions for extremals to reach C_4. (Previously, we heuristically developed (74) and had not shown that (78) was a necessary condition.)

Let us define

$$w=w(p_2(t=T))=\frac{a_1(b_1p_2(t=T)+b_2p)}{p_2(t=T)(a_1b_1-a_2b_2)},$$

which we may write as

$$w=w(Q)=\frac{1}{(R-1)}\left\{R-\frac{\delta}{Q}\right\}, \tag{61}$$

where

$$Q=(-p_2(t=T))/q. \tag{62}$$

Then recalling (29) (which was a necessary condition of optimality), we have that for $\delta < Q \leqslant 1$,

$$1<w(Q)\leqslant\frac{R-\delta}{R-1}=z. \tag{63}$$

Now, $w(Q)$ is a strictly increasing function of Q and $w(Q=\delta)=1$. Also

$$w=\cosh\sqrt{a_2b_2}(T-t_2), \tag{64}$$

and

$$z=\cosh\sqrt{a_2b_2}\tau_1=\cosh\sqrt{a_2b_2}\tau_1(C_5^S).$$

Since $\cosh x$ is a strictly increasing function, (63) and (64) readily yield

$$T-t_2\leqslant\tau_1=\tau_1(C_5^S). \tag{65}$$

Now, the extremal control is

$$\phi^*(t)=\begin{cases}1 & \text{for } 0\leqslant t\leqslant t_2,\\ 0 & \text{for } t_2\leqslant t\leqslant T,\end{cases} \tag{66}$$

where t_2 is the smallest t, such that

$$a_1b_1y^2(t_2)=R\{(b_1x_1(t_2)+b_2x_2^0)^2-(b_1x_1(t_2))^2\}. \tag{67}$$

In [28] we showed that

$$b_1x_1(t_2) = b_2x_2^0(R-1) + \sqrt{s^2 + R(R-1)(b_2x_2^0)^2 - a_1b_1y_0^2}. \tag{68}$$

Clearly we must have $b_1x_1(t_2) \leq b_1x_1^0$, and this may be combined with (68) to yield

$$b_2x_2^0(R-1) \leq b_1x_1^0. \tag{69}$$

A backwards integration of the state equations with the use of (66) yields that $t_2 \leq t \leq T$

$$b_2x_2(t) = b_1x_1(t_2)\{\cosh \sqrt{a_2b_2}\,(T-t) - 1\}, \tag{70}$$

so that

$$b_2x_2(t_2) = b_2x_2^0 = b_1x_1(t_2)\{\cosh \sqrt{a_2b_2}(T-t_2) - 1\}. \tag{71}$$

Then (68) and (71) may be combined to yield

$$T - t_2 = \frac{1}{\sqrt{a_2b_2}} \cosh^{-1}\left\{\frac{Rb_2x_2^0 + \sqrt{s^2 + R(R-1)(b_2x_2^0)^2 - a_1b_1y_0^2}}{(R-1)b_2x_2^0 + \sqrt{s^2 + R(R-1)(b_2x_2^0)^2 - a_1\,b_1y_0^2}}\right\}. \tag{72}$$

Recalling that

$$\tau_1 = \tau_1(C_5^S) = \frac{1}{\sqrt{a_2b_2}} \cosh^{-1}\left(\frac{R-\delta}{R-1}\right), \tag{73}$$

we may combine (65) and (72) to obtain

$$a_1b_1y_0^2 \leq s^2 + A(b_2x_2^0)^2. \tag{74}$$

Now, since $\phi^*(t) = 1$ for $0 \leq t \leq t_2$, we have by the generalized square law (7)

$$a_1b_1y_0^2 - a_1b_1y^2(t_2) = s^2 - (b_1x_1(t_2) + b_2x_2^0)^2. \tag{75}$$

We may combine (67) and (75) to obtain

$$a_1b_1y_0^2 = (R-1)\{(b_1x_1(t_2) + b_2x_2^0)^2 - s^2\} + R\{(b_1x_1^0)^2 - (b_1x_1(t_2))^2\} + R\{s^2 - (b_1x_1^0)^2\}. \tag{76}$$

We will prove as Proposition 1 that

$$R\{(b_1x_1^0)^2 - (b_1x_1(t_2))^2\} - (R-1)\{s^2 - (b_1x_1(t_2) + b_2x_2^0)^2\} \geq 0, \tag{77}$$

so that (76) and (77) yield

$$a_1 b_1 y_0^2 \geq R\{s^2 - (b_1 x_1^0)^2\}. \tag{78}$$

It remains to prove the following proposition

PROPOSITION 1: A necessary and (with appropriate additional assumptions) sufficient condition for (77) to be true is that $b_2 x_2^0 (R-1) \leq b_1 x_1^0$.

PROOF: We prove necessity only (sufficiency follows by reversing the chain of arguments).

It is readily shown that (77) may be rearranged to yield

$$(b_1 x_1^0)^2 - (b_1 x_1(t_2))^2 \geq 2(R-1) b_2 x_2^0 \{b_1 x_1^0 - b_1 x_1(t_2)\}. \tag{79}$$

Now (79) may be factored and (assuming that $b_1 x_1^0 > b_1 x_1(t_2)$) a common term cancelled on both sides to yield

$$b_1 x_1^0 + b_1 x_1(t_2) \geq 2(R-1) b_2 x_2^0. \tag{80}$$

Combining (68) with (80), we obtain

$$\sqrt{s^2 + R(R-1)(b_2 x_2^0)^2 - a_1 b_1 y_0^2} \geq b_2 x_2^0 (R-1) - b_1 x_1^0, \tag{81}$$

and the necessity of the proposition is proved. Q.E.D.

The time t_2 in (66) may be explicitly determined by solving the equation (see Table VII)

$$y(t=t_2) = y_0 \cosh \sqrt{a_1 b_1}\, t_2 - \frac{s}{\sqrt{a_1 b_1}} \sinh \sqrt{a_1 b_1}\, t_2, \tag{82}$$

and this yields

$$t_2 = \frac{1}{\sqrt{a_1 b_1}} \ln \left\{ \frac{y(t_2) - \sqrt{y^2(t_2) - y_0^2 + s^2/(a_1 b_1)}}{y_0 - s/\sqrt{a_1 b_1}} \right\}. \tag{83}$$

Since the quantity in brackets must be ≥ 1 (for all cases, i.e., $a_1 b_1 y_0^2 >, =, < s^2$), we must have

$$y_0 \geq y(t_2) > 0. \tag{84}$$

Now (67) and (68) may be combined to yield

$$y^2(t_2) = \frac{x_2^0}{a_2} \left\{ (2R-1) b_2 x_2^0 + 2\sqrt{s^2 + R(R-1)(b_2 x_2^0)^2 - a_1 b_1 y_0^2} \right\}, \tag{85}$$

so that when this is combined with (84), we obtain

$$\left\{\frac{a_2b_2y_0^2}{b_2x_2^0}-(2R-1)b_2x_2^0\right\}\geq 2\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}. \tag{86}$$

Hence, we must have

$$a_1b_1y_0^2\geq R(2R-1)(b_2x_2^0)^2. \tag{87}$$

If (87) holds, then squaring both sides of (86) again leads to (78).

Moreover, (69) and (78) yield (87) as Proposition 2 shows.

PROPOSITION 2: Assume that $(R-1)b_2x_2^0\leq b_1x_1^0$. Then

$$a_1b_1y_0^2\geq R\{s^2-(b_1x_1^0)^2\} \text{ implies that}$$

$$a_1b_1y_0^2\geq R(2R-1)(b_2x_2^0)^2.$$

The proof to Proposition 2 is straight forward and omitted.

11. DISCUSSION

Let us first consider the structure of the optimal time-sequential fire-distribution policy for the Isbell-Marlow problem (1) before we deduce some important military principles regarding target selection and the valuation of combat resources. The essential *characteristics of the optimal fire-distribution policy*, denoted as ϕ^*, may be summarized as follows:

(1)* ϕ^* is always 0 or 1 (except for at most one point in time),

(2) parameters on which the optimal policy depends are

(a) whether Y wins or loses,

(b) $R=a_1b_1/(a_2b_2)$,

(c) $\delta=a_1p/(a_2q)$.

There are some important military interpretations of the above: (A) a_ib_i is a measure of strategic value of firing at X_i (rate of destruction of X_i's kill capability against Y), and (B) a_1p is a measure of short-run return to Y from firing at X_1 at end of battle (rate of destruction of X_1 value at end of battle).

A significant aspect of the optimal fire-distribution policy (expressed as a closed-loop control) is that it depends on the force levels and not on time. This result is remarkable because the maximum principle does not directly involve the state variables (i.e., the force levels) when the Hamiltonian is maximized for $x_1, x_2 > 0$. Furthermore, the optimal policy for Y may be different depending on whether he wins or loses. Assuming that $R=a_1b_1/(a_2b_2)>1$, then *if* Y *is going to win*, $\phi^*=1$ for $x_1>0$. *If* Y *is going to lose*, then the solution depends on another parameter $(\delta=a_1p/(a_2q))$, and may be very complicated to express as a closed-loop policy. *When* Y *loses*, the general features of Y's optimal

*This is a consequence of Y causing attrition to X_i at a rate proportional to only the number of firers. It is not true in general (see [31]).

fire-distribution policy may be described as follows. Let $p=k(1+\gamma)b_1$ and $q=kb_2$, where γ is a parameter which reflects whether the surviving X_1 forces are valued more ($\gamma>0$) or less ($\gamma<0$) than in direct proportion to their kill capability. Then $p=q(b_1/b_2)(1+\gamma)$, and we have $\gamma=-1+\delta/R$. Moreover, we have (1) $\gamma=0$ means that surviving enemy weapon systems are valued in direct proportion to their kill capability; (2) for $\gamma\geq-(1-1/R)$, simple solution: $\phi^*=1$ for $x_1>0$; (3) for $-(1-1/R)>\gamma\geq-\sqrt{1-1/R}$, it is complicated to obtain solution;† and to (4) $-\sqrt{1-1/R}>\gamma\geq 1$, it is *very* complicated to obtain solution. In the latter two cases, ϕ^* may be initially 1 and then changes to 0 later. When this change occurs is the complicated part.

Let us now discuss what important military principles may be deduced from the solution to the Isbell-Marlow problem. *Firstly*, from the fact that ϕ^* is always 0 or 1, we have a quantitative justification of one of the most significant and oft-quoted of Napoleon Bonaparte's sayings (see p. 117 of [18])—"The principles of war are the same as those of a siege; fire must be concentrated at one point." *Secondly*, from the fact that when Y wins (or when he loses with $\delta\geq 1$) fire is always concentrated on the available enemy target type with largest a_ib_i, we have a quantitative justification of the military principle of attacking "those dangerous enemy targets against which one's fire is effective." *Thirdly*, we have a motivation for valuing enemy target types in direct proportion to their kill capability (fire effectiveness) from the fact that the optimal policy is both intuitively appealing and also very simple in this case. The Howes and Thrall concept of "ideal" linear weights [12] is an extension of this idea to cases of heterogeneous forces on both sides. Thus, we have a motivation for Howes and Thrall's important military valuation methodology. *Fourthly*, in battle a commander must use his judgment to ascertain to what ends the course of battle can be steered so that he may devise his strategy accordingly. Computationally this means that to solve such a problem one must know to which extremal end states* the system can be steered (i.e., what force levels are required to drive the system to a target set, such that appropriate necessary conditions of optimality are satisfied at the end). In other words, it turns out that considerations "in the large" dominate obtaining the optimal policy in such problems.

Let us try to put into perspective what we have learned from our re-examination of the Isbell-Marlow problem. *First of all*, we have concluded that the theory of SVIC's is absolutely essential for the treatment of force-level constraints in such time-sequential combat problems. Many of the results in the pioneering 1956 work of Isbell and Marlow [14] are unsupported by such analysis (which uses theory developed since then). When more than two target types are considered (i.e., n versus one combat), the theory of SVIC's must be used to determine the order in which it is optimal to annihilate target types. As we explained in [31], the theory of SVIC's must be used for extensions of (1) in which force replacements and variable attrition-rate coefficients are considered. *Secondly*, the algebraic method (see [28]) that we have developed for the synthesis of optimal policies appears to us to be the only way of treating such problems.

Our re-examination of the Isbell-Marlow problem has been particularly fruitful, since it revealed the dominated payoff for $0\leq\delta<R-\sqrt{R(R-1)}$, i.e., $P^0\epsilon D(C_5)\cap D(C_5^s)$ implies that $P_5>P_5^s$. We had previously observed such a phenomenon in the supporting weapon system game of H. K. Weiss [37] (see [29]). At that time, we could not relate such a dominated payoff in a differential game to any cor-

†It should be noted that for $R>1$ we have $0<1-1/R<1$.

*The author has developed theoretical results along this line (see [33]).

responding phenomenon in the simpler one-sided (optimal control) case. Additionally, we discovered that local conditions of optimality (i.e., the maximum principle) are, by themselves, inadequate to determine the optimal fire-distribution policy when there are no replacements due to the "absorbing" boundary of the state space. (See discussion of optimal paths to C_1 for $\delta < 1$ in Section 7 (see also [33]).) This, unfortunately, will be true for all such problems.

Finally, we would like to point out that our work here lays a firm foundation for considering two-sided problems. It should be noted that lack of a theory of SVIC's is a gap in the current theory of differential games (particularly as applied to tactical allocation problems in the Lanchester theory of combat).

APPENDIX A. THEOREMS ON THE VALUE OF THE RETURN FUNCTIONAL FOR $R-\sqrt{R(R-1)}<\delta<1$

The purpose of this appendix is to state and prove Theorems A_1, A_2, and A_3.

THEOREM A1: For $R-\sqrt{R(R-1)}<\delta<1$,

$$P_4(P^0) > P_1(P^0) \qquad \text{for all } P^0 \in \{D(C_1) \cap D(C_4)\}.$$

This theorem is a direct consequence of the definitions of $D(C_1)$ and $D(C_4)$ (see Table I) and Lemmas A1.1 and A1.2, which we now state and prove. [It should be noted that a necessary condition for $\{D(C_1) \cap D(C_4)\}$ to be nonempty is that there exist points satisfying

$$s^2-(b_2x_2^0)^2 < a_1b_1y_0^2 \leq s^2+A(b_2x_2^0)^2, \tag{A.1}$$

so that we must have $-1<A$. Moreover, $A<0$ for $R-\sqrt{R(R-1)}<\delta$. By using (5) and other definitions, it may be shown that $-1<A<0$ for

$$R-\sqrt{R(R-1)}<\delta<R-\left(\frac{R-1}{R}\right)\{1+\sqrt{R(R-1)}\}<1. \tag{A.2}$$

(A.2), then, is a necessary condition for $\{D(C_1) \cap D(C_4)\}$ to be nonempty for $R-\sqrt{R(R-1)}<\delta$.]

LEMMA A1.1: Let $a_1b_1y_0^2=s^2+A(b_2x_2^0)^2$. Then $P_4>P_1$ for

$$R-\sqrt{R(R-1)}<\delta<1 \qquad \text{and} \qquad P_4=P_1 \qquad \text{for } \delta=R-\sqrt{R(R-1)}.$$

PROOF: By using the extremal controls for extremals leading to C_1 and C_4, it is readily computed that (see [28])

$$P_1=\left(\frac{-q}{b_2R}\right)\sqrt{R}\,\sqrt{s^2+(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}, \tag{A.3}$$

and

(A.4) $$P_4=\left(\frac{-q}{b_2R}\right)\delta\{(R-1)(b_2x_2^0)+\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}\}.$$

For $a_1b_1y_0^2=s^2+A(b_2x_2^0)^2$, we have

(A.5) $$(P_4-P_1)=\left(\frac{qx_2^0}{R}\right)\{\sqrt{R}\sqrt{-A+R-1}-\delta(R-1)-\delta\sqrt{-A+R(R-1)}\}.$$

By using (5), it is easily shown that

(A.6) $$-A+R(R-1)=\left(\frac{\delta}{z-1}\right)^2,$$

and

(A.7) $$-A+R-1=\frac{2\delta-1}{(z-1)^2}.$$

Next, we show that $2\delta-1>0$ for $R-\sqrt{R(R-1)}<\delta<1$. Now, $2\delta-1$ is a strictly increasing function of δ which is positive for $\delta=R-\sqrt{R(R-1)}$, since it is easily shown that $(2R-1)>2\sqrt{R(R-1)}$.

Substituting (A.6) and (A.7) into (A.5), we have that $P_4>P_1$, if and only if the following inequality holds

(A.8) $$\frac{\sqrt{R}\sqrt{2\delta-1}}{(z-1)}>\delta(R-1)+\frac{\delta^2}{(z-1)}.$$

By using the fact that $(z-1)=(1-\delta)/(R-1)$, some algebraic manipulation yields that (A.8) is true if and only if

(A.9) $$R(2\delta-1)>\delta^2.$$

Now, let us consider $F(\delta)$ defined by

(A.10) $$F(\delta)=R(2\delta-1)-\delta^2.$$

$F(\delta)$ is a strictly increasing function of δ, since $F'(\delta)=2(R-\delta)>0$ for all $R>1>\delta$ with $F(\delta=R-\sqrt{R(R-1)})=0$. Hence, $F(\delta)>0$ for $R-\sqrt{R(R-1)}<\delta<1$, and the lemma is proved.

Q.E.D.

We also have the following lemma.

LEMMA A1.2: For $R>1>\delta$, we have $\frac{\partial}{\partial y_0}(P_1-P_4)>0$.

PROOF: It is readily computed by using (A.3) and (A.4) that

A. 11)
$$\frac{\partial}{\partial y_0}(P_1-P_4)=\frac{2a_1b_1y_0q}{b_2R}\left\{\frac{\sqrt{R}}{\sqrt{s^2+(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}}-\frac{\delta}{\sqrt{s^2+R(R-1)(b_2x_2^0)^2-a_1b_1y_0^2}}\right\},$$

hence follows the lemma. Q.E.D.

THEOREM A2: For $R-\sqrt{R(R-1)}<\delta<1$,

$$P_5(P^0)>P_1(P^0) \qquad \text{for all } P^0\epsilon\{D(C_1)\cap D(C_5)\}.$$

his theorem is a direct consequence of the definitions of $D(C_1)$ and $D(C_5)$ (see Table I) and Lemmas 2.1 and A2.2, which we now state and prove.

LEMMA A2.1: Let $a_1b_1y_0^2=Rs^2(1-1/z^2)$. Then $P_5>P_1$ for

$$R-\sqrt{R(R-1)}<\delta<1 \qquad \text{and} \qquad P_5=P_1 \qquad \text{for } \delta=R-\sqrt{R(R-1)}.$$

LEMMA A2.2: For $R>1$ and $b_1x_1^0>0$, we have $\frac{\partial}{\partial y_0}(P_1-P_5)>0$. The above lemmas follow rom arguments given on p. 550 of [28].

THEOREM A3: For $R-\sqrt{R(R-1)}<\delta<1$,

$$P_5^S(P^0)>P_1(P^0) \qquad \text{for all } P^0\epsilon\{D(C_1)\cap D(C_5^S)\}.$$

his theorem is a direct consequence of the definitions of $D(C_1)$ and $D(C_5^S)$ (see Table I) and the llowing lemma. (It should be recalled that $A<0$ for $R-\sqrt{R(R-1)}<\delta<1$.)

LEMMA A3.1: For $R-\sqrt{R(R-1)}<\delta<1$ and $a_1b_1y_0^2<s^2$, we have $P_5^S>P_1$.

PROOF: For $R-\sqrt{R(R-1)}<\delta<1$, use of the extremal control for extremal leading to C_5^S yields

A.12)
$$P_5^S=\left(\frac{-q}{b_2R}\right)\{z(R-1)(b_2x_2^0)+\sqrt{\{R-z^2(R-1)\}(s^2-a_1b_1y_0^2)}\}.$$

Using (A.3) and (A.12), we have that $P_5^S>P_1$ if and only if

A.13)
$$z(R-1)(b_2x_2^0)+\sqrt{\{R-z^2(R-1)\}(s^2-a_1b_1y_0^2)}<\sqrt{R\{s^2+(R-1)(b_2x_2^0)^2-a_1b_1y_0^2\}}.$$

The condition that $R-\sqrt{R(R-1)}<\delta$ is equivalent to $R>z^2(R-1)$ by use of the definition of . Now for $s^2>a_1b_1y_0^2$ and $R>z^2(R-1)$, we have

A.14)
$$\{z\sqrt{s^2-a_1b_1y_0^2}-(b_2x_2^0)\sqrt{R-z^2(R-1)}\}^2>0,$$

hence follows (A.13) and the lemma by some algebraic manipulation. Q.E.D.

APPENDIX B. RELATIONSHIP BETWEEN THE RESULTS OF BRYSON, DENHAM, AND DREYFUS AND THOSE OF SPEYER AND BRYSON

Let us consider the problem (denoted as "Problem I.")

Problem I.
$$\underset{u(t)}{\text{maximize}}\ G(T, x_i(T)) + \int_0^T L(t, x_i, u)dt,$$

subject to:
$$\frac{dx_k}{dt} = f_k(t, x_i, u) \qquad \text{for } k = 1, \ldots, n,$$

$u(t)$ is unrestricted scalar control variable, $C(t, x_i(t)) \leqslant 0$ for all $t \epsilon [0, T]$ (scalar inequality constraint on state variables), where we assume that all functions are smooth enough to insure the existence of all partial derivatives required. We further assume that the first time derivative of $C(t, x_i(t))$ explicitly contains u and that $(\dot{C})_u(t, x_i(t)) \neq 0$ along an optimal trajectory.

Bryson, Denham, and Dreyfus [5] (formally) adjoin the first time derivative of the state constraint to the return functional with a multiplier (μ) for boundary segments to insure feasibility and obtained the following necessary conditions (see also [15])

(B.1)
$$\frac{\partial H_B}{\partial u} = L_u + \sum_{k=1}^{n} \lambda_k \frac{\partial f_k}{\partial u} - \mu(\dot{C})_u = 0,$$

(B.2)
$$\frac{d\lambda_k}{dt} = -\frac{\partial H_B}{\partial x_k} = -L_{x_k} - \sum_{j=1}^{n} \lambda_j \frac{\partial f_j}{\partial x_k} + \mu(\dot{C})_{x_k} \qquad \text{for } k = 1, \ldots, n,$$

where

(B.3)
$$H_B = L + \sum_{k=1}^{n} \lambda_k f_k - \mu \frac{dC}{dt},$$

and

(B.4)
$$\mu(t) \begin{cases} = 0 & \text{for } C < 0, \\ \geqslant 0 & \text{for } C = 0. \end{cases}$$

Additionally, we must have [23] (see also [32]) $\dot{\mu}(t) \leqslant 0$ when $C = 0$ for a finite interval of time. At exists (at t_l) from constrained subarcs we have $\mu(t_l^-) = \mu(t_l^+) = \mu(t_l) = 0$.

Speyer and Bryson [27] (formally) adjoin the state constraint itself to the return frunctional with a multiplier (η) and obtained the following necessary conditions (see also [15])

(B.5)
$$\frac{\partial H_S}{\partial u} = L_u + \sum_{k=1}^{n} p_k \frac{\partial f_k}{\partial u} - \eta C_u = 0,$$

(B.6)
$$\frac{dp_k}{dt} = -\frac{\partial H_S}{\partial x_k} = -L_{x_k} - \sum_{j=1}^{n} p_j \frac{\partial f_j}{\partial x_k} + \eta C_{x_k} \qquad \text{for } k = 1, \ldots, n,$$

where

(B.7) $$H_S = L + \sum_{k=1}^{n} p_k f_k - \eta C,$$

and

(B.8) $$\eta(t) \begin{cases} = 0 & \text{for } C < 0, \\ \geq 0 & \text{for } C = 0. \end{cases}$$

At entrances and exists from constrained subarcs [19] we have

(B.9) $$p_k(t_e^-) = p_k(t_e^+) \qquad \text{for } k = 1, \ldots, n,$$

and

(B.10) $$H_S(t_e^-) = H_S(t_e^+).$$

We now show that (B.5) and (B.6) may be considered to be a rearrangement of (B.1) and (B.2). Let us rewrite (B.3) as

(B.11) $$H_B = L + \sum_{k=1}^{n} \left(\lambda_k - \mu \frac{\partial C}{\partial x_k}\right) f_k - \mu \frac{\partial C}{\partial t}.$$

Also, (B.2) may be rewritten as

$$\frac{d\lambda_k}{dt} = -L_{x_k} - \sum_{j=1}^{n} \left(\lambda_j - \mu \frac{\partial C}{\partial x_j}\right) \frac{\partial f_j}{\partial x_k} + \mu \frac{d}{dt}\left(\frac{\partial C}{\partial x_k}\right),$$

or

(B.12) $$\frac{d}{dt}\left(\lambda_k - \mu \frac{\partial C}{\partial x_k}\right) = -L_{x_k} - \sum_{j=1}^{n} \left(\lambda_j - \mu \frac{\partial C}{\partial x_j}\right) \frac{\partial f_j}{\partial x_k} - \dot{\mu} \frac{\partial C}{\partial x_k}.$$

Considering (B.6) and (B.12), we may make the identifications

(B.13) $$p_k = \lambda_k - \mu \frac{\partial C}{\partial x_k} \qquad \text{for } k = 1, \ldots, n,$$

and

(B.14) $$\eta = -\dot{\mu},$$

since we also have $\partial H_B/\partial u = \partial H_S/\partial u$ when (B.13) holds.* Additionally, (B.4), (B.9), and (B.13) show that the adjoint variables (i.e., λ's) of Bryson et al., [5] may be discontinuous at entrances to contrained subarcs but are continuous at exits where $\mu(t_l) = 0$.

Finally, we give an interpretation of Speyer's multiplier η as the rate of marginal return of having

*It should be noted that $\partial C/\partial t$ does not explicitly depend on u.

the state constraint active. It is convenient to consider the constraint $C(t, x_i) \leq b$. Then, (at least formally (see [15])) the augmented return functional is given by

$$J^* = G + \int_0^T \left\{ L + \sum_{k=1}^{n} p_k (f_k - \dot{x}_k) - \eta(C - b) \right\} dt,$$

where J^* indicates that we are using the optimal control u^*. Differentiating with respect to the parameter b and setting $b = 0$ to obtain Problem I., we obtain

$$\left. \frac{\partial J^*}{\partial b} \right|_{b=0} = \int_0^T \eta(t) dt. \tag{B.15}$$

The result (B.15) is the control theory analogue of the well-known finite dimensional optimization result for Lagrange multipliers (see [1]).

REFERENCES

[1] Afriat, S., "Theory of Maxima and the Method of Lagrange," SIAM J. Appl. Math. *20*, 343–357 (1971).

[2] Antosiewicz, H., "Analytic Study of War Games," Nav. Res. Log. Quart. *2*, 181–208 (1955).

[3] Boltyanskii, V., R. Gamkrelidze, and L. Pontryagin, "On the Theory of Optimal Processes," *Reports of the Academy of Sciences of the USSR 110*, 7–10 (1956) (translation reprinted in R. Bellman and R. Kalaba (Eds.), *Mathematical Trends in Control Theory* (Dover, New York, 1964)).

[4] Bryson, A. and Y. C. Ho, *Applied Optimal Control* (Blaisdell Publishing Company, Waltham, Massachusetts, 1969).

[5] Bryson, A., W. Denham, and S. Dreyfus, "Optimal Programming Problems with Inequality Constraints I: Necessary Conditions for Extremal Solutions," AIAA Journal *1*, 2544–2550 (1963).

[6] Chewning, W., and J. Taylor, "On the Existence of an Optimal Policy for a Continuous Time, Deterministic Inventory Model," submitted for publication.

[7] Davis, B. and D. J. Elzinga, "The Solution of an Optimal Control Problem in Financial Modeling," Operations Res. *19*, 1419–1433 (1971).

[8] Giamboni, L., A. Mengel, and R. Dishington, "Simplified Model of a Symmetric Tactical Air War," The RAND Corporation, RM–711 (Aug. 1951).

[9] Harris, K. and L. Wegner, "Tactical Airpower in NATO Contingencies: A Joint Air-Battle/Ground-Battle Model (TALLY/TOTEM)," The RAND Corporation, R–1194–PR (May 1974).

[10] Hestenes, M., "A General Problem in the Calculus of Variations with Applications to Paths of Least Time," RM–100, The RAND Corporation (1949).

[11] Ho, Y. C., "Differential Games, Dynamic Optimization, and Generalized Control Theory," J. Opt. Th. Appl. *6*, 179–209 (1970).

[12] Howes, D. and R. Thrall, "A Theory of Ideal Linear Weights for Heterogeneous Combat Forces," Nav. Res. Log. Quart. *20*, 645–659 (1973).

[13] Isaacs, R., *Differential Games* (John Wiley, New York, 1965).

[14] Isbell, J. and W. Marlow, "Methods of Mathematical Tactics," *Logistics Papers*, No. 14, The George Washington University Research Project (Sept. 1956).

[15] Jacobson, D., M. Lele, and J. Speyer, "New Necessary Conditions of Optimality for Control Problems with State-Variable Inequality Constraints," J. Math. Anal. Appl. *35*, 255–284 (1971).

[16] Lansdowne, Z., G. Dantzig, R. Harvey, and R. McKnight, "Development of an Algorithm to Solve Multi-Stage Games," Control Analysis Corporation, Palo Alto, California (May 1973).

[17] Lee, E., and L. Markus, *Foundations of Optimal Control Theory* (John Wiley & Sons, Inc., New York, 1967).

[18] Liddell Hart, B. H., *Strategy* (Frederick A. Praeger, Inc., New York, 1967) (Second Revised Edition).

[19] McIntyre, J., and B. Paiewonsky, "On Optimal Control with Bounded State Variables," in *Advances in Control Systems*, Leondes (Ed.) (Academic Press, New York, 1967), vol. 5, pp. 389–419.

[20] McNicholas, R., and F. Crane, "Guide to Fire Support Mix Evaluation Techniques, Volume I: The Guide and Appendices A and B," Stanford Research Institute, Menlo Park, California (Mar. 1973).

[21] Neustadt, L. W., "An Abstract Variational Theory with Applications to a Broad Class of Optimization Problems. II. Applications," SIAM J. Control *5*, 90–137 (1967).

[22] Neustadt, L. W., "A General Theory of Extremals," J. Comp. System Sci. *3*, 57–92 (1969).

[23] Pontryagin, L., V. Boltyanskii, R. Gamkrelidze, and E. Mischenko, *The Mathematical Theory of Optimal Processes* (Interscience, New York, 1962).

[24] "Proceedings of the Conference on the State-of-the-Art of Mathematics in Combat Models," to appear.

[25] Pugh, G., "Theory of Measures of Effectiveness for General-Purpose Military Forces: Part II. Lagrange Dynamic Programming in Time-Sequential Combat Games," Operations Res. *21*, 886–906 (1973).

[26] Pugh, G., and J. Mayberry, "Analysis of General-Purpose Force Procurement," Lambda Corporation, Paper 59 (Aug. 1971).

[27] Speyer, J., and A. Bryson, "Optimal Programming Problems with a Bounded State Space," AIAA Journal *6*, 1488–1491 (1968).

[28] Taylor, J., "On the Isbell and Marlow Fire Programming Problem," Nav. Res. Log. Quart. *19*, 539–556 (1972).

[29] Taylor, J., "Some Differential Games of Tactical Interest," Operations Res. *22*, 304–317 (1974).

[30] Taylor, J., "Solving Lanchester-Type Equations for 'Modern Warfare' with Variable Coefficients," Operations Res. *22*, 756–770 (1974).

[31] Taylor, J., "Target Selection in Lanchester Combat: Heterogeneous Forces and Time-Dependent Attrition-Rate Coefficients," Nav. Res. Log. Quart. *21*, 683–704 (1974).

[32] Taylor, J., "Comments on a Multiplier Condition for Problems with State Variable Inequality Constraints," IEEE Trans. on Automatic Control, Vol. AC–17, 743–744 (1972).

[33] Taylor, J., "On Boundary Conditions for Adjoint Variables in Problems with State Variable Inequality Constraints," IEEE Trans. on Automatic Control, Vol. AC–19, 450–452 (1974).

[34] Taylor, J., "Survey on the Optimal Control of Lanchester-Type Attrition Processes," to appear

in Proceedings of the Conference on the State-of-the-Art of Mathematics in Combat Models (also Naval Postgraduate School Technical Report No. NPS55Tw74031, March 1974).

[35] Taylor, J., "On Continuous Time, Deterministic Inventory Models with Linear Production Rate Costs," Cahiers du Centre d'Études Recherche Opérationnelle *15*, 475–486 (1973).

[36] USAF Assistant Chief of Staff, Studies and Analysis, "Methodology for Use in Measuring the Effectiveness of General Purpose Forces, SABER GRAND (ALPHA)," (Mar. 1971).

[37] Weiss, H., "Some Differential Games of Tactical Interest and the Value of a Supporting Weapon System," Operations Res. *7*, 180–196 (1959).

EMPIRICAL BAYES SINGLE SAMPLING PLANS FOR SPECIFIED POSTERIOR CONSUMER AND PRODUCER RISKS*

Harry F. Martz

Los Alamos Scientific Laboratory
Los Alamos, New Mexico

ABSTRACT

Empirical Bayes single-sample acceptance sampling procedures are derived. These procedures assume that the lot fraction defective varies randomly according to a completely unknown and unspecified prior distribution. The unknown prior density function is estimated based on sampling results from previous lots. A procedure is developed for obtaining single-sampling plans that achieve specified posterior consumer and producer risks. The procedures are illustrated for a real-data example.

1. INTRODUCTION

Consider a sequence of $m > 1$ lots submitted for single sampling inspection by attributes in which the binomial is used to approximate the hypergeometric distribution. In the ith lot of size N_i units the number of defectives x_i in a random sample of size $n_i << N_i$ is observed and recorded. The ith lot fraction defective p_i is assumed to be the value of a random variable P. We will sometimes write $x = x_m$, $n = n_m$, and $p = p_m$. For a given single sampling plan (n, c), a lot is accepted when $x \leq c$, the acceptance number, and rejected otherwise. It is further assumed that (i) x_i, conditional on p_i, are independent, and (ii) p_i are independent realizations of P according to a completely unknown and unspecified prior (incoming lot quality) distribution $G(p)$. The past lot history $(x_1, n_1; x_2, n_2; \ldots; x_m, n_m)$ is used in a supplementary manner which by-passes the necessity to specify $G(p)$ and identifies the entire situation as one of Empirical Bayes. A brief introduction to Empirical Bayes is given by Krutchkoff [6] and Martz [10]. A thorough description is provided by Maritz [8].

Recent attention has been directed toward the assumption of a beta prior distribution, the natural conjugate prior distribution for the binomial parameter (Raiffa and Schlaifer [15]). Weiler [18] has shown that the effect of assuming a beta distribution, when in fact the true prior distribution is not of the beta family, is ofttimes negligible. On the other hand, MacFarland [7] has shown that in many cases the assumption of a beta prior distribution leads to undesirable results. The assumption of a beta prior distribution is less easily justified whenever lot fraction defective appears to be concentrated over a range much smaller than [0. 1] (Weiler [18]). A real-data example of such a situation is presented in Section 5. Also, quality control personnel are sometimes reluctant to assume a specified prior distribution, such as a beta, which cannot be easily verified. The Empirical Bayes technique avoids this psychological stigma, provided a suitable past lot history exists. The procedure is applicable without regard for the form of the prior lot quality distribution.

Following along the lines of Weiler [18], the following quantities are of interest:

(1) the unconditional probability $P(P \leq p)$ that a randomly selected lot will have a fraction de-

*Supported by ONR contract N00014-73-A-0434-0001 when the author was at Texas Tech University.

fective P not exceeding p. (This probability characterizes the incoming lot quality of lots supplied by the producer and will be referred to as the Prior Distribution (PD) Function.

(2) the posterior probability $P(P \leq p|x)$ that P does not exceed p in a lot in which sampling has produced x defectives. (This probability characterizes the quality of an inspected lot.)

(3) the posterior probability $P(P \leq p|A)$ that P does not exceed p in accepted lots (denoted by A). (This probability characterizes the quality of lots accepted by the consumer under the plan (n, c) and is of particular interest to the consumer. It is referred to as the Posterior Consumer's Operating Characteristic (PCOC) Curve.)

(4) the posterior probability $P(P \leq p|R)$ that P does not exceed p in rejected lots (denoted by R). (This probability characterizes the quality of lots rejected under a given plan and is of interest to the producer. It is designated as the Posterior Producer's Operating Characteristic (PPOC) Curve.)

(5) the marginal (unconditional) probability $P(X=x)$ that a sample from a randomly selected lot will contain exactly x defectives. (This probability indicates to the consumer the long range sampling results when buying from a given producer.)

(6) the marginal probability, $P(A)$, that a randomly selected lot will be accepted. (This probability represents the long range fraction of submitted lots that will be accepted under a given plan and is of particular interest to the producer and indirect interest to the consumer.)

(7) the mean $E(p|x)$ lot fraction defective in lots in which sampling has produced x defectives. (This posterior mean estimates the lot fraction defective in inspected lots and is of interest to the consumer.)

(8) the mean $E(p|A)$ lot fraction defective in accepted lots. (This posterior mean estimates the expected lot fraction defective in accepted lots and is of particular interest to the consumer.)

Empirical Bayes estimates of these quantities will be considered in Sections 2 and 3, and their performances in Section 4. An Empirical Bayes procedure for obtaining single-sampling plans which satisfy specified posterior consumer's and producer's risks will be presented in Section 5. These procedures will be illustrated by means of a real-data example in Section 6. Finally, some conclusions are presented in the last section.

2. AN EMPIRICAL PRIOR DISTRIBUTION

Parzen [14] provides a class of kernel estimators which can be used in a somewhat modified manner to estimate the prior lot fraction defective density function $f(p)$. A similar approach was used by Bennett and Martz [1], [2], although in a different application. Martz and Lian [13] have also taken such an approach for estimating the parameter of a binomial distribution. The estimator chosen here for estimating $f(p)$ is given by

$$f(p|\mathbf{y}) = \frac{1}{Nh} \sum_{i=i}^{m} K_i \left[\frac{p - y_i}{h} \right], \qquad 0 \leq p \leq 1, \tag{1}$$

where

$$h = h(m) = \begin{cases} W s_m \left(\sum_{i=1}^{m} \delta_i \right)^{-1/5} & (s_m > 0) \\ 2 \left(\sum_{i=i}^{m} \delta_i \right)^{-1/5} / n_m & (s_m = 0), \end{cases} \tag{2}$$

and

$$(3)\qquad K_i(t)=\begin{cases} w_i n_i(1-|t|) & (|t|\le 1,\ y_i-h\ge 0,\ y_i+h\le 1)\\ 2w_i n_i h^2(1-|t|)/\{2h^2-(y_i-h)^2\} & (|t|\le 1,\ y_i-h<0)\\ 2w_i n_i h^2(1-|t|)/\{2h^2-(1-y_i-h)^2\} & (|t|\le 1,\ y_i+h>1). \end{cases}$$

Also, we have defined $\mathbf{y}=(y_1, y_2, \ldots, y_m)$, where $y_i=x_i/n_i$.

In (2)

$$(4)\qquad s_m^2=\Sigma w_i n_i(y_i-\bar{y})^2/N,\ \bar{y}=\Sigma w_i n_i y_i/N,$$

where $N=\Sigma w_i n_i$. Also in (2), $\delta_i=1$ or 0 according to whether $w_i>0$ or $w_i=0$, respectively.

Here w_i is a non-negative weight applied to the data from the ith lot and is free to be selected by the user. Typical reasons for the desirability of such a weight might be that (i) the inspection results for certain atypical lots are not to be included in the calculation of (1) (in this case a zero weight would be assigned to such lots); (ii) lots inspected in the distant past should be weighted less than lots inspected in the recent past (here weights should be assigned based on some function of inspection order). Following along these lines, several different weighting schemes are developed by Martz [11], [12]. Throughout this paper $w_i=1$ for all i.

The constant W appearing in (2) is a non-negative user supplied "window weight." For those situations in which the underlying prior density function is not believed to be multimodal with sharp peaks, this weight should be set equal to one. Otherwise, this weight will generally need to be significantly less than one (for an example see Martz [11], [12]). For the remainder of this paper W will be set equal to one.

Thus the empirical prior density function in (1) consists of the weighted average of m triangular smoothing functions each centered about y_i with total base width $2h$, height $w_i n_i$, and area $w_i n_i h$. It is also noted that if y_i is such that $y_i-h<0$ or $y_i+h>1$, the area of the corresponding triangle is normalized so that its area remains equal to $w_i n_i h$. A transformation of y_i could also be considered (Bennett and Martz [1]). However, in order to keep the estimator as simple as possible, this will not be done here although a slight gain in performance is anticipated as a result of the transformation. Suppose we further assume the existance of the first two moments of the prior distribution. Let μ and σ^2 denote the true (unknown) prior mean and variance of P. It is straightforward to prove the following:

THEOREM: If $s_m>0$ and $w_i=1$, $y_i-h\ge 0$, $y_i+h\le 1$ for all i, then, for the empirical prior density estimator in (1),

$$\mathrm{E}[\mathrm{E}(P|\mathbf{y})]=\mu$$

and

$$E[\mathrm{Var}\ (P\mid \mathbf{y})]=[1+m^{-2/5}/6]\left[\sigma^2\left(\frac{N^2-mN-M+N}{N^2}\right)+\left(\frac{1+m}{N}\right)\mu(1-\mu)\right],$$

where $M=\Sigma n_i^2$.

It is observed that, although the expected mean of (1) agrees with the prior mean, the expected variance of (1) somewhat overestimates the prior variance. However, it is noted that if $n_i\to\infty$, for all i,

and $m \to \infty$, then $\lim E[\text{Var}(P|\mathbf{y})] = \sigma^2$. Thus we have an asymptotic second order match between (1) and $f(p)$.

3. SOME EMPIRICAL BAYES SINGLE-SAMPLING QUANTITIES

Throughout this section, Empirical Bayes estimates of the quantities listed in Section 1 will be designated as being conditional on $\mathbf{y}$. According to the sampling situation described in Section 1, the conditional probability of exactly x defectives in a random sample of size n is given by

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n. \tag{5}$$

First consider the Empirical Bayes estimate of the marginal probability of exactly x defectives in a random sample of size n from the mth lot. We have that

$$P(X = x|\mathbf{y}) = \int_0^1 f(x|p) f(p|\mathbf{y}) dp. \tag{6}$$

Substituting (5) and (1) into (6), integrating, and simplifying yields

$$P(X = x|\mathbf{y}) = \sum_{i=1}^{m} H_i(y_i) Q_i(x, n; 1)/\{Nh^2(n+1)(n+2)\}, \tag{7}$$

where

$$H_i(y_i) = \begin{cases} w_i n_i & (y_i - h \geq 0 \text{ and } y_i + h \leq 1) \\ 2w_i n_i h^2/\{2h^2 - (y_i - h)^2\} & (y_i - h < 0) \\ 2w_i n_i h^2/\{2h^2 - (1 - y_i - h)^2\} & (y_i + h > 1) \end{cases} \tag{8}$$

and where we have defined

$$\begin{aligned} Q_i(x, n; a) = {} & 2(x+a) I_{y_i}(x+a+1, n-x+1) - (x+a) I_{A_i}(x+a+1, n-x+1) \\ & - (x+a) I_{B_i}(x+a+1, n-x+1) + (h+y_i)(n+a+1) I_{B_i}(x+a, n-x+1) \\ & - (h-y_i)(n+a+1) I_{A_i}(x+a, n-x+1) - 2y_i(n+a+1) I_{y_i}(x+a, n-x+1). \end{aligned} \tag{9}$$

Also in (9),

$$A_i \equiv \max(0, y_i - h), \; B_i \equiv \min(1, y_i + h), \tag{10}$$

and $I_z(a, b)$ is the standard incomplete beta function defined by

$$I_z(a, b) = \frac{1}{B(a, b)} \int_0^z t^{a-1} (1-t)^{b-1} dt. \tag{11}$$

By means of (6), whenever sampling inspection is conducted under the plan (n, c), the Empirical Bayes estimate of the marginal probability of lot acceptance becomes

$$P(A \mid \mathbf{y}) \equiv P(X \leq c \mid \mathbf{y}) = \sum_{x=0}^{c} P(X = x \mid \mathbf{y}) = \sum_{i=1}^{m} H_i(y_i) \sum_{x=0}^{c} Q_i(x, n; 1)/\{Nh^2(n+1)(n+2)\}. \tag{12}$$

Now let us consider an Empirical Bayes estimator for the posterior probability $P(P \leq p \mid x)$ discussed in Section 1. We have

$$P(P \leq p \mid x, \mathbf{y}) = \frac{\int_0^p f(x \mid p) f(p \mid \mathbf{y}) dp}{P(X=x \mid \mathbf{y})} \qquad \{P(X=x \mid \mathbf{y}) \neq 0\} \tag{13}$$

$$= 0 \qquad \{P(X=x \mid \mathbf{y}) = 0\}.$$

Substituting (5), (1), and (6) into (13), integrating, and simplifying yields

$$P(P \leq p \mid x, \mathbf{y}) = \frac{\sum_{i=1}^{m} H_i(y_i) b(p, A_i) R_i(p, x, n)}{\sum_{i=1}^{m} H_i(y_i) Q_i(x, n; 1)} \qquad \{P(X=x \mid \mathbf{y}) \neq 0\} \tag{14}$$

$$= 0 \qquad \{P(X=x \mid \mathbf{y}) = 0\},$$

where

$$b(u, v) = \begin{cases} 1 & (u > v) \\ 0 & (u \leq v), \end{cases} \tag{15}$$

and where we have defined

$$\begin{aligned} R_i(p, x, n) &= (n+2)(h-y_i)[I_{C_i}(x+1, n-x+1) - I_{A_i}(x+1, n-x+1)] \\ &+ (x+1)[I_{C_i}(x+2, n-x+1) - I_{A_i}(x+2, n-x+1)] \\ &+ (n+2) b(p, y_i)\{(h+y_i)[I_{D_i}(x+1, n-x+1) \\ &- I_{y_i}(x+1, n-x+1)] \\ &- (x+1)[I_{D_i}(x+2, n-x+1) \\ &- I_{y_i}(x+2, n-x+1)]\}. \end{aligned} \tag{16}$$

Also, in (16), we have defined

$$C_i \equiv \min\ (p, y_i), \qquad D_i \equiv \min\ (p, B_i). \tag{17}$$

It is easily shown that if $p \geq \max_i (y_i + h)$, then (14) is equal to 1.

The posterior mean of the density function corresponding to (14) is the Empirical Bayes estimator (under squared error loss) for the lot fraction defective in lots for which x defectives are obtained from a sample of size n and becomes

$$E(p \mid x, \mathbf{y}) = \frac{x+1}{n+3} \cdot \frac{\sum_{i=1}^{m} H_i(y_i) Q_i(x, n; 2)}{\sum_{i=1}^{m} H_i(y_i) Q_i(x, n; 1)}. \tag{18}$$

The risk in using this estimator has been considered by Martz and Lian [13].

Let us now consider Empirical Bayes estimators for the Posterior Consumer's and Producer's Operating Characteristic (PCOC and PPOC) curves. For the Empirical Bayes PCOC (EBPCOC) curve we obtain

$$P(P \leq p|A, \mathbf{y}) = \frac{\sum_{i=1}^{m} H_i(y_i) b(p, A_i) \sum_{x=0}^{c} R_i(p, x, n)}{\sum_{i=1}^{m} H_i(y_i) \sum_{x=0}^{c} Q_i(x, n; 1)} \qquad \{P(A|\mathbf{y}) \neq 0\} \tag{19}$$

$$= 0 \qquad \{P(A|\mathbf{y}) = 0\}.$$

Again it is observed that if $p \geq \max_i (y_i + h)$ then this probability is equal to 1 provided that $P(A|\mathbf{y}) \neq 0$. It is also noted that (19) is a cumulative probability distribution over the range $[\max\{0, \min_i(y_i - h)\}, \min\{1, \max_i (y_i + h)\}]$.

Similarly, the Empirical Bayes PPOC (EBPPOC) curve becomes

$$P(P \leq p|R, \mathbf{y}) = \frac{0.5(n+1)(n+2) \sum_{i=1}^{m} H_i(y_i) b(p, A_i) S_i(p) - \sum_{i=1}^{m} H_i(y_i) b(p, A_i) \sum_{x=0}^{c} R_i(p, x, n)}{Nh^2(n+1)(n+2) - \sum_{i=1}^{m} H_i(y_i) \sum_{x=0}^{c} Q_i(x, n; 1)} \tag{20}$$

$$= 0 \qquad \{P(A|\mathbf{y}) \neq 1\}$$

$$\{P(A|\mathbf{y}) = 1\},$$

where we have defined

$$S_i(p) = 2(y_i - h)(A_i - C_i) + (C_i - A_i)(C_i + A_i) + b(p, y_i)(p - y_i)(2h + y_i - p). \tag{21}$$

Likewise, if $p \geq \max_i (y_i + h)$ and $P(A|\mathbf{y}) \neq 1$, then this probability is also equal to 1. As before, the EBPPOC curve is a cumulative distribution over the same range as the EBPCOC curve.

The Empirical Bayes estimator for the expected lot fraction defective in lots accepted under the plan (n, c) becomes

$$E(p \mid A, \mathbf{y}) = \frac{\sum_{i=1}^{m} H_i(y_i) \sum_{x=0}^{c} (x+1) Q_i(x, n; 2)}{(n+3) \sum_{i=1}^{m} H_i(y_i) \sum_{x=0}^{c} Q_i(x, n; 1)}. \tag{22}$$

Finally, the prior distribution corresponding to (1), the Empirical Bayes Prior Distribution (EBPD) Function, may be directly obtained by integration of (1) as

$$P(P \leq p \mid \mathbf{y}) = \sum_{i=1}^{m} H_i(y_i) b(p, A_i) S_i(p) / \{2Nh^2\}. \tag{23}$$

It is remarked that, although all of these Empirical Bayes quantities look formidable, they may be easily and efficiently evaluated on any computer. Martz [11] describes a basic FORTRAN IV program which has been written for calculating most of these Empirical Bayes quantities in practice. This source deck and related documentation are available from the author upon request.

4. PERFORMANCE EVALUATION

In order to assess the statistical performance characteristics of the estimators presented in Section 3, a limited Monte Carlo simulation was conducted. An upper limit on m was set at 150 (although m as large as 205 will be reported in Section 6). For convenience, the binomial sample size n_i was considered to be constant and ranged from 2 to 500. A piecewise continuous linear family of prior distributions $f(p)$ was constructed which consisted of up to eight adjoining straightline segments over the range, or a subset of the range, [0, 1]. Such a linear family permitted the (true) exact Bayes quantities corresponding to those in Section 3 to be determined for purposes of comparison with the corresponding Empirical Bayes quantities. For certain cases, the corresponding quantities were also computed under the assumption that the prior distribution was a member of the beta family, although in fact the true prior distribution was a piecewise linear function as described above. Such estimators are referred to as Beta-Bayes quantities.

The complete simulation results reported in Martz [9] indicate that generally for unimodal symmetric or moderately skewed priors the Empirical Bayes and Beta-Bayes estimators yield results which are in good agreement with each other and both agree quite well with the corresponding (true) Bayes quantities. However, this is not the case for highly skewed (such as U-shaped) priors. In this case, a greater variation in the performance of the Empirical Bayes and Beta-Bayes quantities was observed with the Empirical Bayes procedure yielding somewhat of an improvement over the Beta-Bayes method. This improvement was especially pronounced when estimating the (true) Bayes PCOC and PPOC curves.

First and second sample moments of the estimators in Section 3 were also computed by averaging over the number of replications used in the simulation (usually 100). The results indicated that the Empirical Bayes estimators quite satisfactorily approximate the corresponding true (but unknown in practice) Bayes quantities for a wide variety of prior distributions, binomial sample sizes n, and values of m. The results of a typical comparison are presented in Figures 1, 2, and 3 for the case of a uniform prior distribution [0, 0.04], $n=80$, $m=25$, and sampling plan (80, 1). In addition, Figure 1 contains the (true) Bayes $P(A)$ and $E(p|A)$ values as well as the corresponding Empirical Bayes estimates. The curves and values presented are averages over the 100 replications and are thus estimates of the first moments of the Empirical Bayes estimators. The sample means of the Empirical estimators are observed to be in good agreement with the (true) Bayes quantities, except for the upper range of the PD Function and PPOC Curve. However, the upper range of the PPOC Curve is not generally used in deriving suitable single-sampling plans (see Sec. 5). The corresponding sample standard deviations are reasonably small and are given in Martz [9]. It was found that as m increases the standard deviations decreased rapidly with the largest gradient occuring between $m=2$ and $m=25$. By $m=25$, the Empirical Bayes estimators are smooth and stable in their mean.

5. SAMPLING PLAN DETERMINATION

Let us now consider a procedure for finding the single-sampling plan (if one exists) that attains

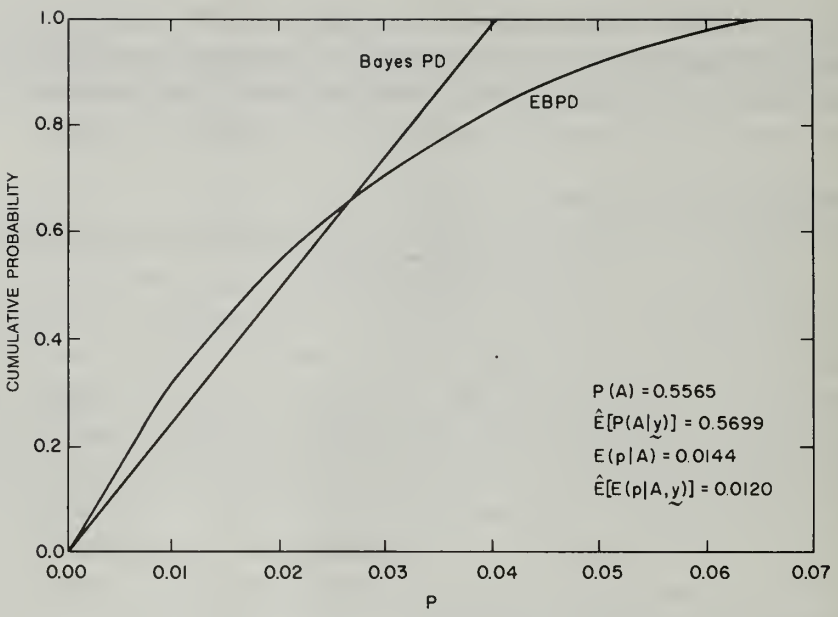

FIGURE 1. Average PD functions for a uniform lot quality distribution, $n = 80$, $m = 25$, and 100 replications.

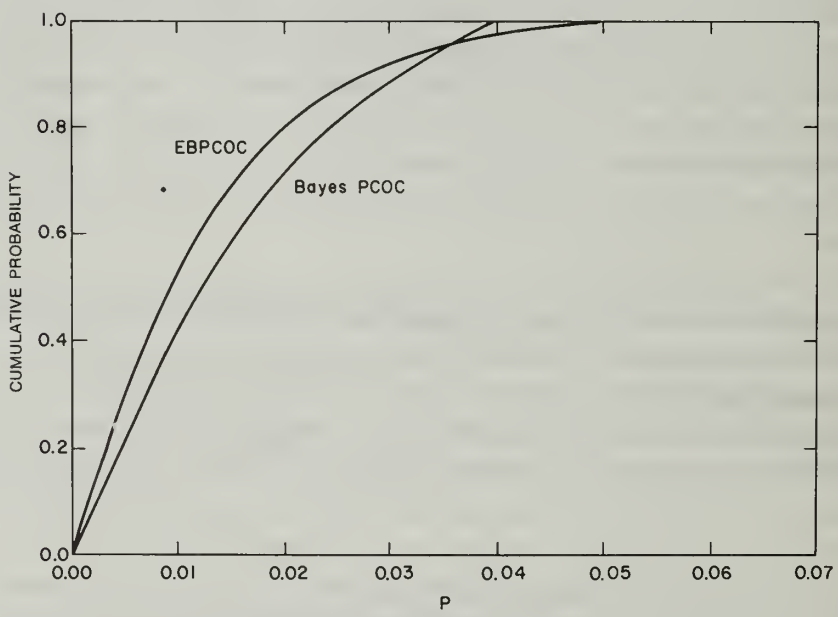

FIGURE 2. Average PCOC curves under the plan (80, 1) for a uniform lot quality distribution, $n = 80$, $m = 25$, and 100 replications.

specified Empirical Bayes posterior consumer's and producer's risks. Cameron [4] presents a method for determining plans which attain specified (classical) consumer and producer risks. Analogous Bayesian procedures have been considered by Hald [5] and Schafer [16, 17].

Similar to Schafer, let us define the following two Empirical Bayes posterior risks when using the plan (n, c):

(24) Empirical Bayes Posterior Consumer's Risk $\equiv \beta(n, c) \equiv P(P > p_2 | A, \mathbf{y})$,

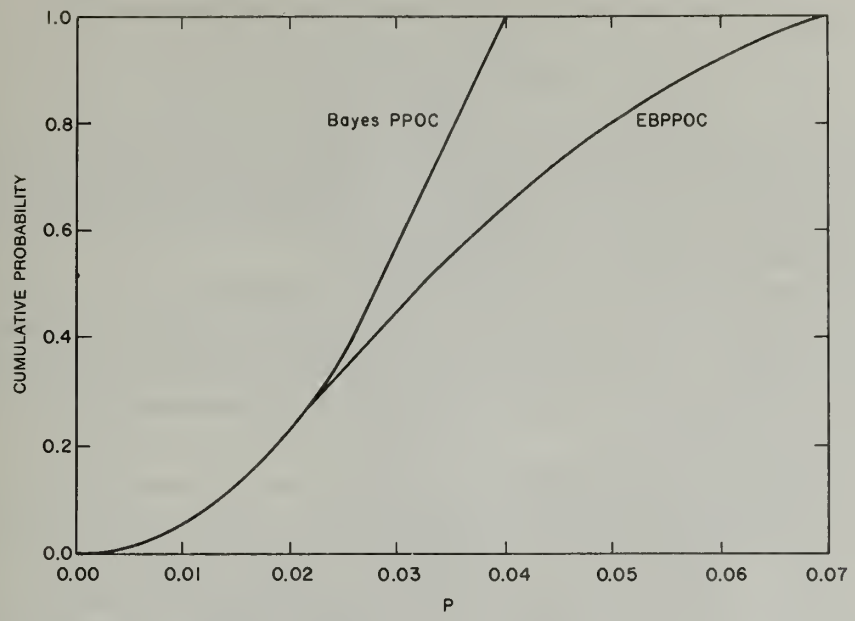

FIGURE 3. Average PPOC curves under the plan (80, 1) for a uniform lot quality distribution, $n = 80$, $m = 25$, and 100 replications.

where p_2 is a particular specified value of p such that the consumer requires a high probability that P does not exceed this value in accepted lots. The value p_2 is analogous to the classical Lot Tolerance Percent Defective (LTPD), expressed as a decimal fraction, although the interpretation of this quantity is different. Similarly we define the

(25) Empirical Bayes Posterior Producer's Risk $\equiv \alpha(n, c) \equiv P(P < p_1 | R, \mathrm{y})$,

where p_1 is a particular specified value of p such that the producer requires a low probability that P does not exceed this value in rejected lots. Also, p_1 is analogous to the classical Acceptable Quality Level (AQL) expressed as a decimal fraction.

It is desired by the consumer and producer that a single-sampling plan (n, c) be found (if one exists) which satisfies the conditions

(26) $$\beta(n, c) \leq \beta \ (\beta \text{ Small})$$

(27) $$\alpha(n, c) \leq \alpha \ (\alpha \text{ Small}),$$

where β and α are maximum allowable values of the specified risks. Also, as in the case of classical procedures, p_2 must be greater than p_1. Plans are frequently desired for α, $\beta = 0.10$, 0.05, or 0.01, although other values are sometimes specified.

Before proceeding to discuss the procedure used to determine the required plan (provided a plan exists), an important observation is made. The producer does not really desire any inspection at all and would be well pleased if the consumer purchased his lots without inspection. This situation would occur if a sufficient number of lots of excellent quality have previously been submitted by the producer.

For example, suppose for the past lot history that $p_2 \geq \max_i (y_i + h)$. Then from (19),

$$P(P > p_2 | A, \mathbf{y}) = 0 \leq \beta$$

for any β and for any sampling plan. Hence, in particular, the Empirical Bayes posterior consumer's risk is satisfied for the plan $n = 0$. Obviously the producer will not insist upon sampling if the consumer is willing to purchase his unsampled material. Hence, the optimal procedure which satisfies both the consumer and the producer is to purchase the unsampled lot.

Now let us consider a procedure for determining the required plan. Generally, there exists more than one single-sampling plan that satisfies the inequalities (26) and (27). As in Schafer [16, 17] we denote the "optimal plan" as that plan which satisfies (26) and (27), and which has the smallest value of n within the class of all plans that satisfy (26) and (27). If this plan is not unique, then the one with the largest $P(A | \mathbf{y})$ is selected.

It is desired to locate the optimal plan for a range of values of n between 0 and some preselected upper bound n_u. The value of n_u is arbitrarily selected and, for example, may be (i) selected such that it is a certain fraction (such as 0.10) of the lot size; (ii) set equal to the lot size; or (iii) that sample size value currently being used for lots of this size. When in doubt, a good rule-of-thumb is to set n_u equal to the lot size. Thus there exist a total of $(n_u+1)\ (n_u+2)/2$ possible sampling plans (n, c), where $0 \leq n \leq n_u$ and $0 \leq c \leq n$, which must be searched for the optimal plan, provided that such a plan exists which may or may not be the case as discussed later.

The search procedure to be described depends heavily upon the following observations which have been made concerning the Empirical Bayes single-sampling quantities of Section 3.

For fixed n

$\beta\ (n, c)$ increases as c increases

$\alpha\ (n, c)$ decreases as c increases

$P\ (A | \mathbf{y})$ increases as c increases

For fixed c

$\beta\ (n, c)$ decreases as n increases

$\alpha\ (n, c)$ increases as n increases

$P\ (A/\mathbf{y})$ decreases as n increases

Similar observations were also made by Schafer, although in a Bayesian context.

Since the above observations correspond to those of Schafer, the direct search method he employs to find the optimal plan may be used here (Schafer [16]). A computer program flow diagram corresponding to the search technique used here is given in Figure 4. The justification for this search procedure may be found in Schafer [16], p. 18, and is based upon the above observations which, in summary, state that a reduction in the Empirical Bayes posterior consumer's risk is paid for an increase in the consumer's risk and vice versa.

A basic FORTRAN IV program was written to conduct the search outlined in Figure 4 and is documented in Martz [12]. The source deck and documentation are available from the author upon request.

There is one important difference between Schafer's results and those here. Schafer showed that,

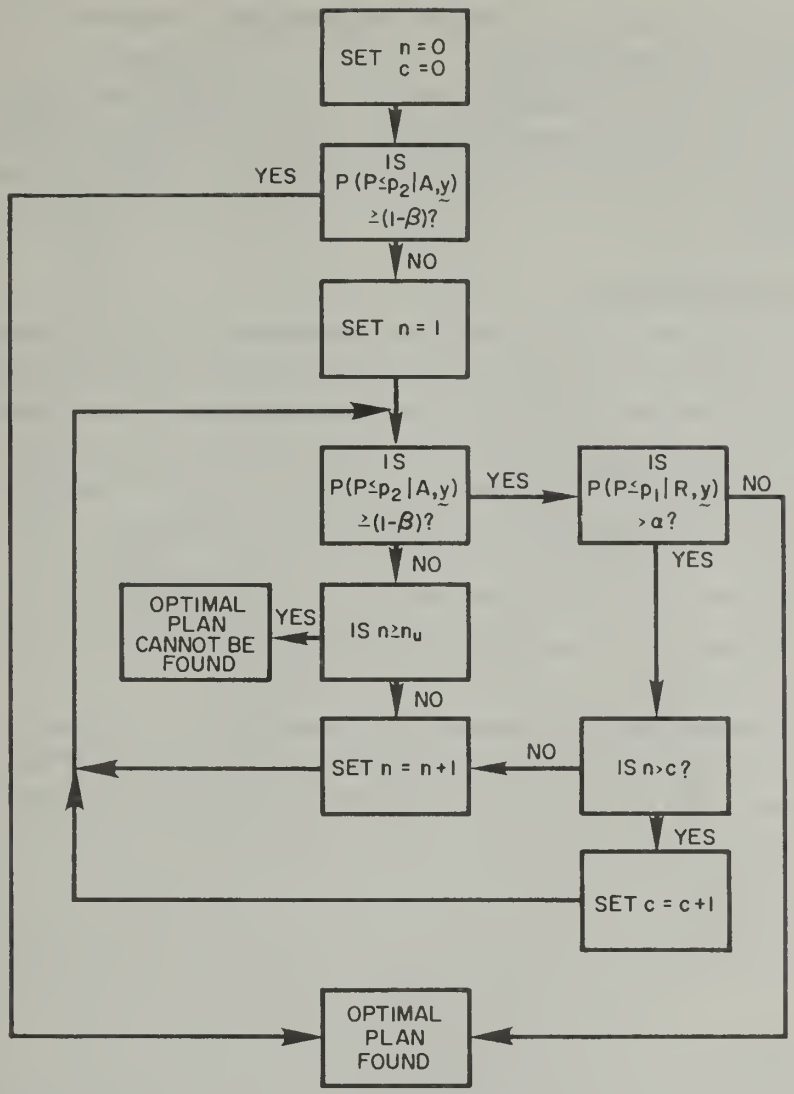

FIGURE 4. Computer program flow diagram for determining the optimal Empirical Bayes single-sampling plan.

within the framework he adopts, an optimal solution always exists. Such is not the case here. Within the grid of possible plans there may not exist an optimal plan due to large tail areas in the empirical prior density function either to the right of p_2 or to the left of p_1 or both. Also, the optimal sample size n, in general, increases as α and/or β decreases. Thus α and/or β may be so small as to force the optimal solution outside the chosen grid.

Suppose that the optimal sampling plan cannot be found within the confines of the grid selected. Then, based on the observed magnitude of the risks capable of being produced within the grid, (i) p_2 should be increased, accordingly; (ii) p_1 should be decreased; (iii) α and/or β should be increased; or (iv) n_u should be increased. The actual measures to be taken depend upon individual preference and upon that risk which is incapable of being satisfied.

It is noted here that if $P(A|\mathbf{y}) = 1$, then $P(P \leq p_2|A, \mathbf{y}) = P(P \leq p_2)$ and, by definition, $P(P \leq p_2|R, \mathbf{y}) \equiv 0$. Such is the case for the plan (0, 0) which is the first plan evaluated according

to Figure 4. Hence, if the plan (0, 0) satisfies the Empirical Bayes consumer's risk then the optimal procedure is to purchase the unsampled lot.

Finally, the optimal sampling plan found here may or may not yield a sample size n which is smaller than the classical sample size. Generally, if prior lot quality has been good, the optimal sample size will be smaller (reduced inspection), while the opposite is true for poor previously observed lot quality (tightened inspection).

6. A REAL-DATA EXAMPLE

J. C. Borg [3], in his unpublished Lehigh University Master's Thesis, presented some real-data which apparently originated from the Bureau of Naval Weapons. The data were obtained under the use of the MIL–STD–105B plan and consist of the observed number of minor, major, and critical defects in successive samples of some vendor produced material. However, only the minor defect data will be considered here. The data in Table 1 consist of the observed frequencies of the number of minor defectives in samples of size 150 from 205 lots each of size 2016 items.

TABLE 1. *Observed Frequencies of Minor Defectives Per Lot*

($n = 150$, $m = 205$)

Number of defectives (x)	Frequency	Number of defectives (x)	Frequency
0	68	9	2
1	45	10	1
2	24	12	1
3	20	13	1
4	8	15	4
5	7	18	1
6	8	20	1
7	10	22	1
8	3		
		Total.....205	

The observed range of the sample lot fraction defective values is [0, 0.15]. It appears extremely unlikely that the prior distribution is concentrated over the entire range [0, 1] unless, of course, the right tail area is very small. Nevertheless, using goodness-of-fit tests, Borg tested the appropriateness of assuming a beta prior distribution for this data. He concluded that the prior is adequately described by the beta distribution. In spite of this, we will illustrate the Empirical Bayes estimators for this data.

For the minor defect sampling inspection an AQL = 2.5 percent was selected and the corresponding sampling plan used in inspecting each of the 205 lots was ($n = 150$, $c = 8$). Normal inspection at Level II was considered. Figure 5 presents five Empirical Bayes quantities for this plan, corresponding to constant binomial sample size $n = 150$, $m = 205$, and for the data in Table 1.

From the EBPPOC curve in Figure 5, we observe that $P(P < 0.025 | R, \mathbf{y}) \cong 0.007$, which is roughly 50-percent smaller than the calculated $\alpha = 0.015$ corresponding to the AQL = 2.5 percent for the classical plan (150,8). It must be kept in mind, however, that the comparison is not strictly a fair one since the interpretation of the producer's risk is different in each case.

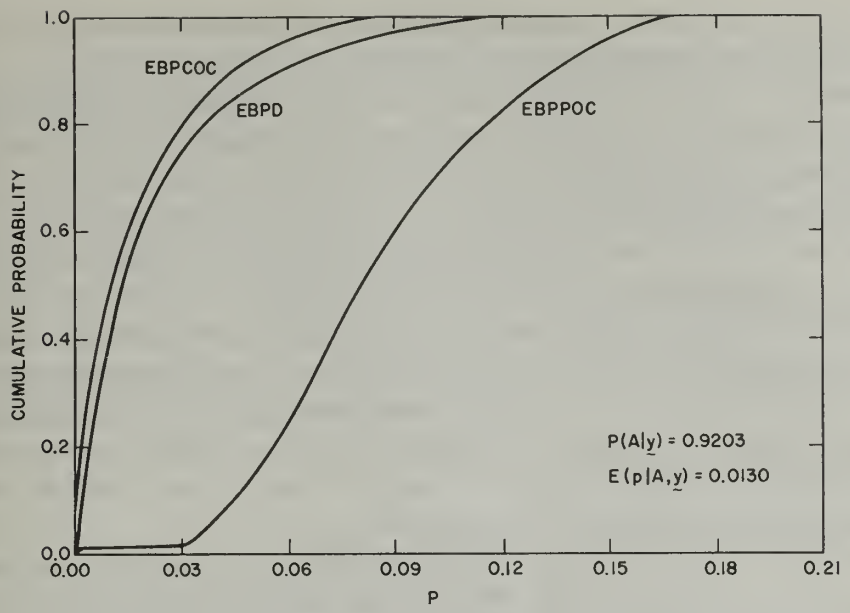

FIGURE 5. Empirical Bayes single-sampling quantities for Borg (1962) minor defect data.

It is also seen from Figure 5 that 92 percent of submitted lots are estimated to be accepted under the plan. From Table 1, it is observed that actually 94 percent of the 205 lots have been accepted. It is also estimated that accepted lots are expected, on the average, to be 1.3 percent defective.

The Beta-Bayes quantities corresponding to those in Figure 5 were also computed. The beta prior distribution parameters were estimated to be $n' = 32.26$ and $r' = 0.56$ (see Weiler [18]), from which $P(A)$ and $E(p|A)$ were estimated to be 0.9245 and 0.0128. These estimates, under the assumption of a beta prior distribution, agree quite closely with those given in Figure 5. The PCOC, PPOC, and PD curves were also found to be in good agreement with those in Figure 5.

Now suppose we desire the plan which satisfies the risk values $\alpha = 5$ percent, $\beta = 10$ percent, $p_1 = 2.5$ percent, and $p_2 = 4.0$ percent for the data in Table 1. That is, we wish to determine the single-sampling plan (if one exists) such that $P(P > 0.04|A, \mathbf{y}) \leq 0.10$ and $P(P < 0.025(R, \mathbf{y}) \leq 0.05$. Table 2 gives several intermediate results as well as the final result. The required plan is found to be

TABLE 2. *The Empirical Bayes Posterior Consumer and Producer Risks for Borg (1962) Minor Defect Data*

($\alpha = 5$ percent, $p_1 = 2.5$ percent, $\beta = 10$ percent, $p_2 = 4.0$ percent)

n	c	$\alpha(n, c)$	$\beta(n, c)$	$P(A\|\mathbf{y})$	n	c	$\alpha(n, c)$	$\beta(n, c)$	$P(A\|\mathbf{y})$
0	0	0.000	0.141	1.000	19	2	0.043	0.123	0.976
1	0	0.328	0.134	0.982	25	2	0.056	0.111	0.960
5	0	0.353	0.110	0.917	31	2	0.070	0.100	0.942
7	0	0.364	0.100	0.889	31	3	0.019	0.120	0.973
7	1	0.105	0.131	0.985	35	3	0.023	0.113	0.965
15	1	0.140	0.109	0.948	40	3	0.028	0.106	0.955
19	1	0.157	0.098	0.927	44	3	0.034	0.099	0.946

($n = 44$, $c = 3$). It is estimated that roughly 95 percent of all submitted lots will be accepted under this plan, and that those lots accepted will, on the average, be 1.5 percent defective.

For comparison, suppose we obtain the classical single-sampling plan with the above risks. The classical plan is found to be ($n = 1235$, $c = 40$). Here, of course, the data in Table 1 are not used in obtaining this plan. For this classical plan, $\alpha(1235,40) \approx 0.049$ and $\beta(1235,40) \approx 0.087$. Although a comparison between these plans is not strictly a fair one as discussed previously; nevertheless, we do observe an enormous decrease in the Empirical Bayes sample size required. For the classical plan ($n = 44$, $c = 3$), we find that $\alpha = (44,3) \cong 0.03$, at $p_1 = 0.025$, which satisfies the producers risk. However, $\beta(44,3) = 0.88$, at $p_2 = 0.04$, which far exceeds the required consumer's risk.

The Empirical Bayes plan for $\alpha = 1.5$, $\beta = 5$, $p_1 = 2.5$, and $p_2 = 4.0$ percent for the data in Table 1, is found to be ($n = 206$, $c = 9$). Here the sample size is larger than the MIL–STD–105B plan ($n = 150$, $c = 8$). However, for this MIL–STD plan the consumer's risk for $p_2 = 0.04$ is computed to be 0.847 which far exceeds 0.05 guaranteed by the Empirical Bayes plan. To achieve equivalent risks, the classical plan would require a much larger sample size than $n = 206$.

It is mentioned here that the Empirical Bayes plan will not yield a smaller sample size in all cases. It will yield a smaller sample size in those cases in which the past lot quality history has been observed to be good, as demonstrated above. The sample size will generally be larger than the classical sample size if the opposite is the case since more protection is desired.

7. CONCLUSIONS

The Empirical Bayes single-sampling estimators considered here have two basic uses. First, they may be used to evaluate the performance of single-sampling procedures in current use by a consumer, such as the MIL–STD–105D plans. The Empirical Bayes estimators can indicate any discrepancies that may exist between what these classical plans are intended to do and what they actually do, based upon the empirical evidence and assuming a Bayesian interpretation for the classical plans. This use is essentially that of a management tool for use in evaluating existing procedures (Martz [11]). In this regard, the Empirical Bayes estimators can be added to an existing quality control program without any discrediting effect on existing sampling procedures.

Second, the Empirical Bayes estimators can be used to obtain single-sampling plans that achieve specified posterior consumer and producer risks (Martz [12]). These sampling plans have several advantages over classical sampling plans. They indicate the situation in which no sampling inspection is required, i.e., $n = 0$. They also automatically revert to reduced or tightened inspection in situations where such a change is indicated. The sample sizes obtained from these plans are generally smaller than the classical sample sizes (reduced inspection) if previous lots have been observed to be of good quality. If the opposite is true, then the sample sizes are generally larger (tightened inspection).

The main advantage of such Empirical Bayes estimators is that the form of the prior distribution does not have to be explicitly identified. This avoids the stigma of identifying a form for the prior which may or may not be realized in practice.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. R. E. Schafer, Hughes Aircraft Co., and a referee for helpful comments regarding this manuscript.

REFERENCES

[1] Bennett, G. K. and H. F. Martz, "A Continuous Empirical Bayes Smoothing Technique," Biometrika *59*, 361–368 (1972).

[2] Bennett, G. K. and H. F. Martz, "An Empirical Bayes Estimator for the Scale Parameter of the Two Parameter Weibull Distribution," Nav. Res. Log. Quart. *20*, 387–393 (Sept. 1973).

[3] Borg, J. C., "Sampling Inspection Where a Continuing Relationship Exists Between Consumer and Producer," Master's Thesis, Lehigh University (1962).

[4] Cameron, J. M., "Tables for Constructing and for Computing the Operating Characteristics of Single-Sampling Plans," Industrial Quality Control *9*, 37–39 (July 1952).

[5] Hald, A., "The Determination of Single-Sampling Attribute Plans with Given Producers and Consumers Risk," Technometrics *9*, 401–415 (1967).

[6] Krutchkoff, R. G., "Empirical Bayes Estimation," The American Statistician *26*, 14–16 (Dec. 1972).

[7] MacFarland, W. J., "Bayes' Equation, Reliability, and Multiple Hypothesis Testing," IEEE Transactions on Reliability (Aug. 1972), Vol. R–21, No. 3.

[8] Maritz, J. S., "Empirical Bayes Methods," *Methuen's Monographs on Applied Probability and Statistics* (Methuen and Co., Ltd, London 1970).

[9] Martz, H. F., "Empirical Bayes Single-Sample Acceptance Sampling," Technical Report No. 1, Department of Industrial Engineering, Texas Tech University, Lubbock, Texas 79409 (Jan. 31, 1974).

[10] Martz, H. F., "Pooling Life Test Data by Means of the Empirical Bayes Method," IEEE Transactions on Reliability, (Apr. 1975), Vol. R–24, No. 2.

[11] Martz, H. F., "EBSPE–A FORTRAN Computer Program for Conducting an Empirical Bayes Posterior Performance Evaluation of Any Single-Sampling by Attributes Sampling Plan," Technical Report No. 4, Department of Industrial Engineering, Texas Tech University, Lubbock, Texas 79409 (Dec. 31, 1974).

[12] Martz, H. F., "EBSSP–A FORTRAN Computer Program for Determining Empirical Bayes Single-Sampling Plans for Specified Posterior Consumer and Producer Risks," Technical Report No. 3, Department of Industrial Engineering, Texas Tech University, Lubbock, Texas 79409 (Dec. 31, 1974).

[13] Martz, H. F. and M. G. Lian, "Empirical Bayes Estimation of the Binomial Parameter," Biometrika *61*, 517–523 (Dec. 1974).

[14] Parzen, E., "On Estimation of a Probability Density Function and Mode," Annals of Math. Stat. *33*, 1065–1076 (1962).

[15] Raiffa, H. and R. Schlaifer, *Applied Statistical Decision Theory*, Harvard University Press, Boston, (1961).

[16] Schafer, R. E., "Development of Bayes Single Sampling Plans by Attributes, Ph.D. Dissertation, Case Western Reserve University (1968).

[17] Schafer, R. E., "Bayes Single Sampling Plans by Attributes Based on the Posterior Risks," Nav. Res. Log. Quart. *14*, 81–88 (Mar. 1967).

[18] Weiler, H., "The Use of Incomplete Beta Functions for Prior Distributions in Binomial Sampling," Technometrics *7*, 335–347 (Aug. 1965).

AN ANALYSIS OF THE LONG RANGE OPERATING CHARACTERISTICS OF THE MIL-STD-105D SAMPLING SCHEME AND SOME SUGGESTED MODIFICATIONS

Gerald G. Brown

Naval Postgraduate School, Monterey, California

and

Herbert C. Rutemiller

California State University, Fullerton

ABSTRACT

An integral part of the MIL-STD-105D scheme for sampling inspection by attributes is the transfer from normal inspection to reduced or tightened inspection when the historical record of inspected lots suggests unusually good or bad quality. The switching rules in MIL-STD-105D have been criticized, especially by Japanese manufacturers, as being too severe when what is defined as acceptable quality material is submitted. This paper examines the long range fraction of lots rejected for several MIL-STD-105D sampling plans by using the MIL-STD-105D switching rules, using a modification suggested by the Japanese Standards Association, and by using a second modification developed by the authors.

The Japanese Standards Association switching rules are more complex than those in MIL-STD-105D. It is demonstrated that they lead to improved long-range properties for Normal-Tightened-Reduced schemes, but to poorer properties for Normal-Tightened schemes.

A simplified set of switching rules is suggested, wherein the "limit numbers" in MIL-STD-105D are eliminated. In comparison to MIL-STD-105D, the simpler rules lead to a lower probability of rejection for good lots and a comparable probability of rejection for bad lots.

1. INTRODUCTION

The sampling scheme designated as MIL-STD-105D [4] has been widely adopted throughout the western world for lot-by-lot sampling inspection by attributes. The sampling plans in this document are indexed by lot size and by a number called the acceptable quality level (AQL). The AQL is specified by the consumer. The AQL is a designated percentage defective which, if met by the supplier, will lead to acceptance of the great majority of lots [4, sec. 4.3]. MIL-STD-105D provides two parameters for each sampling plan, a sample size, n, and acceptance number, c. If c or fewer defectives are found in a random sample of size n, the lot is accepted.

Fundamental to the MIL-STD-105D scheme is the concept of "normal" and "tightened" inspection. Sampling inspection begins with the normal plan, and the user is required to maintain an historical record of the results for each lot inspected. If two out of any five consecutive lots are rejected, a switch to tightened inspection is required. Under tightened inspection, n is usually the same as for normal inspection, but c is reduced, decreasing the probability of acceptance for all incoming quality levels. Normal inspection may be resumed if five consecutive lots are accepted. If tightened inspection continues for 10 consecutive lots, then sampling inspection must be discontinued.

An optional feature of MIL–STD–105D is the provision for "reduced" inspection when the historical quality record is unusually good. A move from normal to reduced inspection requires that the previous 10 lots have been accepted and that the total number of defectives in the 10 lots does not exceed a set of "limit numbers." Some examples are shown in Table 1.* Under reduced inspection, n is substantially decreased, and two numbers, c and r $(> c)$ are supplied. Lots are accepted if the number of defectives is less than r, but normal inspection must be resumed if the number of defectives is more than c.

It is the contention of many critics of MIL–STD–105D that the switching rules are too severe when incoming quality is in the vicinity of the AQL. Furthermore, the switching rules are difficult to learn and administer, especially with respect to the transfer from normal to reduced inspection. The purpose of this paper is to examine some long-range operating characteristics of MIL–STD–105D and two alternative schemes which provide some modifications in the switching rules.

2. THE CONCEPT OF "LONG RANGE" O.C. CURVES

It is clear that tightened, normal, and reduced inspection will each have a different probability of acceptance for a given incoming lot quality. Hence, the long-range proportion of lots accepted will be dependent on the proportion of lots inspected under each plan.

It is traditional in the literature of quality control to examine the performance of an attributes sampling plan under the assumption that, when the process is "in control," a stream of product is being produced with a fixed probability, p, that each item is defective (see Duncan [2, p. 147] or Grant and Leavenworth [3, p. 364]). The value of p for a specific manufacturing process is often well established. We may think of p as a parameter of the in-control process. Even if fraction defective is not a constant from lot to lot in a particular process, the long-range O.C. curve is of considerable value. For example, the vendor may be able to set a value of p as an upper bound of fraction defective for his process, whereupon the long-range O.C. curve provides an upper bound for fraction of lots rejected.

3. METHODS AND ASSUMPTIONS

In a previous paper [1], we have described the theoretical basis for development of a long-range O.C. curve for the MIL–STD–105D scheme. Tightened and reduced inspection are treated as Markov chains, and normal inspection as a semi-Markov chain. If the process percent defective is nonzero, the MIL–STD–105D scheme leads to repeated visits to tightened inspection, and eventually to tightened inspection of 10 consecutive lots and the discontinuation of sampling inspection. We assume that, when this occurs, the next 10 lots will be treated as *rejected* lots and subjected to 100 percent inspection, after which normal sampling inspection may be resumed.

4. PROPOSED MODIFICATIONS

Ohmae [5] described some changes in the switching rules which were claimed to lead to a high probability of acceptance for AQL quality lots with a smaller probability than MIL–STD–105D of invoking tightened inspection for such lots. The Ohmae modifications eventually led to a new set of switching rules, which have been adopted by the Japanese Standards Association. These changes are shown in Table 2.

*See tables at end of paper.

It is apparent that the Ohmae proposals make the switching rules for transfer from normal to tightened inspection more complicated than the original MIL–STD–105D scheme through the introduction of another set of limit numbers, and that the principal effect will be to alter the probability of a transfer to tightened inspection for all quality levels, and to decrease the probability of discontinuing sampling inspection when tightened inspection is adopted.

We suggest a different modification, also shown in Table 2. Limit numbers for transfer to reduced or tightened inspection are completely eliminated. Lots are viewed as a series of discrete subsets of five. Two consecutive subsets with no rejections are required to reach reduced inspection, unless there has been a rejection under normal inspection. Whenever a rejection occurs under normal inspection, the next four lots must be accepted to prevent a transfer to tightened inspection. If these four lots are accepted, then two additional subsets of five lots with no rejections are required to reach reduced inspection.

5. RESULTS

We have examined the long-range fraction of lots accepted under the three schemes for several lot sizes and for four values of AQL—0.4-, 1.0-, 1.5-, and 2.5-percent defective. At each AQL, the incoming quality level was chosen as 0.2 AQL, 0.6 AQL, 0.8 AQL, 1.2 AQL, 2.0 AQL, and 3.0 AQL.

Normal-Tightened Schemes

Table 3 shows the results for a sampling scheme wherein only normal and tightened inspection are employed.

The Japanese modification has the undesirable characteristic of decreasing, for all AQL's investigated, the fraction of lots accepted when material of AQL or better quality is submitted.

The authors' proposed modification leads to virtually the same long range O.C. curve as MIL–STD–105D.

Normal-Tightened-Reduced Schemes

Table 4 shows the results for the three schemes when reduced inspection is included. Referring to the limit numbers (Table 1) for transfer to reduced inspection, it is apparent that the user of these plans has an option regarding reduced inspection when small lots are inspected (as indicated by the asterisk in Table 1). He may interpret Table 1 as forbidding the use of reduced inspection for these cases, or as permitting the inclusion of more than 10 lots to achieve the necessary sample size. In the former case, the results from Table 3 for these lot sizes should be used. Table 4 has been constructed under the assumption that for plans labeled with an asterisk in Table 1, more than 10 lots may be used to obtain sufficient items.

Where reduced inspection is included, Table 4 shows that, in general, the Japanese modifications provide superior long range O.C. curves to those of MIL–STD–105D. For lots of AQL quality or better, the fraction of lots accepted is nearly always higher than MIL–STD–105D, while for poorer quality lots, the fraction accepted becomes lower than MIL–STD–105D.

The authors' suggested modification is usually even better than the Japanese plan with respect to acceptance of lots at AQL quality or better, but the fraction of poorer quality lots accepted is a little higher, usually comparable to that provided by MIL–STD–105D. This simpler approach, without the

complication of limit numbers, appears to satisfy the objective of accepting more high quality lots while retaining the long range operating characteristics of MIL–STD–105D for poorer quality lots.

A subsidiary advantage of the authors' scheme is worth mentioning. The elimination of limit numbers not only simplifies record keeping, but also permits a complete analysis of the normal sampling scheme as a pure Markov chain. Under these conditions, long range O.C. curves may be obtained without costly iterative solution required for semi Markov analyses.

TABLE 1. *Limit numbers from MIL-STD-105D, Table VIII for transfer from normal to reduced inspection. Numbers in body of table are maximum defective in most recent 10 lots*

Number of items in last 10 lots	AQL				Number of items in last 10 lots	AQL			
	0.40	1.0	1.5	2.5		0.40	1.0	1.5	2.5
20–29	*	*	*	*	320–499	*	0	1	4
30–49	*	*	*	*	500–799	0	2	3	7
50–79	*	*	*	*	800–1,249	0	4	7	14
80–129	*	*	*	0	1,250–1,999	2	8	13	24
130–199	*	*	0	0	2,000–3,149	4	14	22	40
200–319	*	0	0	2	3,150–4,999	8	25	38	67

*MIL–STD–105D requires at least 200 items to be inspected, so the number of items in last 10 lots is not sufficient to qualify for reduced inspection. In these instances, more than 10 lots may be used, if desired.

TABLE 2.—*Switching rules for MIL–STD–105D and two suggested modifications*

Switch	MIL–STD–105D	Japanese modification	Authors' proposed modification
Normal to tightened	2 out of any 5 consecutive lots rejected.	a. A lot is rejected *and* b. Total defectives in last 5 lots $\geq$ limit number.*	Rejected lot, followed by a second rejection in the following 4 lots
Tightened to normal	5 consecutive lots accepted.	No change	No change from 105D.
Tightened to "Discontinue"	10 consecutive lots remain on tightened.	Number of cumulative lots rejected under tightened reaches 5.	No change from 105D.
Normal to Reduced	a. 10 consecutive lots accepted and *and* b. Total defectives in the 10 lots $\leq$ limit number. *and* c. Production at a steady rate *and* d. Reduced inspection is deemed desirable.	No change, except that limit numbers are reduced.	a. 10 consecutive lots accepted from the beginning of normal inspection. *or* b. 14 consecutive lots accepted after a rejection during normal inspection c. and d. same as 105D.
Reduced to Normal	Any lot with more than c defectives.	No change in switching rules, except limit number, c, for plans with $r=2$, is altered from 0 to 1.	

*The new limit numbers are the upper 0.5-percent probability points from the Poisson distribution, assuming that AQL quality material has been submitted.

TABLE 3. *Long range operating characteristics of three sampling schemes employing normal and tightened inspection (MIL-STD-105D general inspection level II)*

AQL = 0.4 percent

Lot size code letter	Incoming percent defective	Fractions of lots accepted		
		MIL-STD-105D	Japanese Standard Assn.	Authors' modification
G	0.08	0.974	0.969	0.975
	0.24	0.905	0.895	0.913
	0.32	0.856	0.846	0.870
	0.40	0.802	0.789	0.817
	0.48	0.746	0.729	0.761
	0.80	0.553	0.516	0.559
	1.20	0.403	0.352	0.404
K	0.08	0.995	0.993	0.995
	0.24	0.958	0.938	0.961
	0.32	0.915	0.880	0.926
	0.40	0.845	0.792	0.865
	0.48	0.754	0.686	0.776
	0.80	0.429	0.348	0.433
	1.20	0.233	0.178	0.233
L	0.08	0.999	0.995	0.999
	0.24	0.986	0.951	0.987
	0.32	0.966	0.898	0.971
	0.40	0.925	0.816	0.941
	0.48	0.854	0.716	0.883
	0.80	0.477	0.367	0.487
	1.20	0.237	0.173	0.238
M	0.08	1.000	0.999	1.000
	0.24	0.992	0.978	0.993
	0.32	0.979	0.949	0.980
	0.40	0.949	0.893	0.957
	0.48	0.886	0.801	0.907
	0.80	0.452	0.362	0.458
	1.20	0.193	0.143	0.193
N	0.08	1.000	1.000	1.000
	0.24	0.998	0.985	0.998
	0.32	0.994	0.959	0.994
	0.40	0.981	0.895	0.983
	0.48	0.943	0.778	0.958
	0.80	0.440	0.283	0.450
	1.20	0.133	0.081	0.133

TABLE 3. *Long range operating characteristics of three sampling schemes employing normal and tightened inspection (MIL-STD-105D general inspection level II)* — Continued

AQL = 1.0 Percent

Lot size code letter	Incoming percent defective	Fraction of lots accepted		
		MIL-STD-105D	*Japanese* Standards Assn.	Authors' modification
E	0.2	0.974	0.969	0.974
	0.6	0.904	0.894	0.912
	0.8	0.854	0.845	0.867
	1.0	0.799	0.788	0.814
	1.2	0.743	0.728	0.757
	2.0	0.550	0.514	0.556
	3.0	0.400	0.351	0.402
H	0.2	0.995	0.993	0.995
	0.6	0.958	0.938	0.961
	0.8	0.915	0.880	0.926
	1.0	0.845	0.792	0.865
	1.2	0.754	0.686	0.776
	2.0	0.429	0.348	0.433
	3.0	0.233	0.178	0.233
J	0.2	0.999	0.995	0.999
	0.6	0.986	0.951	0.987
	0.8	0.966	0.898	0.971
	1.0	0.925	0.816	0.941
	1.2	0.854	0.716	0.883
	2.0	0.477	0.367	0.487
	3.0	0.237	0.173	0.238
K	0.2	1.000	0.999	1.000
	0.6	0.993	0.979	0.993
	0.8	0.980	0.950	0.981
	1.0	0.951	0.896	0.958
	1.2	0.890	0.806	0.910
	2.0	0.459	0.368	0.465
	3.0	0.196	0.145	0.196
L	0.2	1.000	1.000	1.000
	0.6	0.998	0.985	0.998
	0.8	0.994	0.959	0.994
	1.0	0.981	0.895	0.983
	1.2	0.943	0.778	0.958
	2.0	0.440	0.283	0.450
	3.0	0.133	0.081	0.133

TABLE 3. *Long range operating characteristics of three sampling schemes employing normal and tightened inspection (MIL-STD-105D general inspection level II)* — Continued

AQL = 1.5 percent

Lot size code letter	Incoming percent defective	Fractions of lots accepted		
		MIL-STD-105D	Japanese Standards Assn.	Authors' modification
D	0.3	0.976	0.970	0.976
	0.9	0.911	0.900	0.920
	1.2	0.866	0.853	0.880
	1.5	0.815	0.799	0.831
	1.8	0.762	0.741	0.777
	3.0	0.571	0.530	0.578
	4.5	0.418	0.364	0.420
G	0.3	0.996	0.993	0.996
	0.9	0.962	0.945	0.965
	1.2	0.925	0.896	0.934
	1.5	0.864	0.820	0.882
	1.8	0.782	0.724	0.803
	3.0	0.460	0.382	0.465
	4.5	0.255	0.197	0.255
H	0.3	0.999	0.996	0.999
	0.9	0.989	0.958	0.989
	1.2	0.973	0.914	0.976
	1.5	0.941	0.845	0.953
	1.8	0.884	0.755	0.919
	3.0	0.527	0.408	0.540
	4.5	0.268	0.196	0.269
J	0.3	1.000	0.999	1.000
	0.9	0.994	0.981	0.994
	1.2	0.982	0.956	0.983
	1.5	0.958	0.910	0.964
	1.8	0.908	0.831	0.925
	3.0	0.495	0.401	0.503
	4.5	0.216	0.161	0.217
K	0.3	1.000	1.000	1.000
	0.9	0.999	0.988	0.999
	1.2	0.995	0.968	0.996
	1.5	0.986	0.921	0.987
	1.8	0.962	0.827	0.970
	3.0	0.515	0.338	0.533
	4.5	0.166	0.101	0.166

TABLE 3. *Long range operating characteristics of three sampling schemes employing normal and tightened inspection (MIL–STD–105D general inspection level II)* — Continued

AQL = 2.5 Percent

Lot size code letter	Incoming percent defective	Fractions of lots accepted		
		MIL–STD–105D	Japanese Standards Assn.	Authors' modification
C	0.5	0.975	0.969	0.975
	1.5	0.907	0.896	0.916
	2.0	0.851	0.847	0.873
	2.5	0.806	0.791	0.821
	3.0	0.750	0.731	0.766
	5.0	0.558	0.518	0.564
	7.5	0.407	0.354	0.408
F	0.5	0.995	0.993	0.995
	1.5	0.958	0.938	0.961
	2.0	0.915	0.880	0.926
	2.5	0.845	0.792	0.865
	3.0	0.754	0.686	0.776
	5.0	0.429	0.348	0.433
	7.5	0.233	0.178	0.233
G	0.5	0.999	0.995	0.999
	1.5	0.986	0.951	0.987
	2.0	0.966	0.898	0.971
	2.5	0.925	0.816	0.941
	3.0	0.854	0.716	0.883
	5.0	0.477	0.367	0.487
	7.5	0.237	0.173	0.238
H	0.5	1.000	0.999	1.000
	1.5	0.993	0.979	0.993
	2.0	0.980	0.950	0.981
	2.5	0.951	0.896	0.958
	3.0	0.890	0.806	0.920
	5.0	0.459	0.368	0.465
	7.5	0.196	0.145	0.196
J	0.5	1.000	1.000	1.000
	1.5	0.998	0.985	0.998
	2.0	0.994	0.959	0.994
	2.5	0.981	0.895	0.983
	3.0	0.943	0.778	0.958
	5.0	0.440	0.283	0.450
	7.5	0.133	0.081	0.133

TABLE 4. *Long range operating characteristics of three sampling schemes employing normal, tightened and reduced inspection (MIL-STD-105D general inspection level II)*

AQL = 0.4 percent

Lot size code letter	Incoming percent defective	Fractions of lots accepted		
		MIL-STD-105D	Japanese Standards Assn.	Authors' modification
G*	0.08	0.987	0.987	0.988
	0.24	0.935	0.942	0.944
	0.32	0.889	0.901	0.902
	0.40	0.832	0.846	0.852
	0.48	0.770	0.780	0.792
	0.80	0.559	0.531	0.571
	1.20	0.403	0.355	0.407
K	0.08	0.998	0.999	0.998
	0.24	0.970	0.983	0.972
	0.32	0.930	0.947	0.935
	0.40	0.859	0.871	0.870
	0.48	0.764	0.755	0.778
	0.80	0.430	0.355	0.434
	1.20	0.233	0.178	0.233
L	0.08	1.000	1.000	1.000
	0.24	0.996	0.991	0.997
	0.32	0.984	0.972	0.989
	0.40	0.947	0.933	0.968
	0.48	0.871	0.852	0.916
	0.80	0.478	0.384	0.498
	1.20	0.237	0.173	0.238
M	0.08	1.000	1.000	1.000
	0.24	0.997	0.994	0.998
	0.32	0.986	0.981	0.991
	0.40	0.957	0.955	0.971
	0.48	0.892	0.984	0.919
	0.80	0.452	0.370	0.460
	1.20	0.193	0.143	0.193
N	0.08	1.000	1.000	1.000
	0.24	1.000	0.999	1.000
	0.32	0.996	0.994	0.998
	0.40	0.983	0.981	0.993
	0.48	0.945	0.938	0.970
	0.80	0.440	0.294	0.641
	1.20	0.133	0.081	0.133

*16 lots required to meet minimum sample assignment of Table VIII for AQL = 0.4, 1.0, 2.5. 17 lots required for AQL = 1.5.

TABLE 4. *Long range operating characteristics of three sampling schemes employing normal, tightened and reduced inspection (MIL-STD-105D general inspection level II)* — Continued

AQL = 1.0 Percent

Lot size code letter	Incoming percent defective	Fractions of lots accepted		
		MIL-STD-105D	Japanese Standards Assn.	Authors' modification
E*	0.2	0.987	0.988	0.988
	0.6	0.935	0.942	0.944
	0.8	0.889	0.901	0.904
	1.0	0.831	0.845	0.855
	1.2	0.768	0.778	0.791
	2.0	0.556	0.529	0.568
	3.0	0.401	0.353	0.404
H	0.2	0.998	0.999	0.998
	0.6	0.970	0.983	0.972
	0.8	0.930	0.947	0.935
	1.0	0.859	0.871	0.870
	1.2	0.764	0.755	0.778
	2.0	0.430	0.355	0.434
	3.0	0.233	0.178	0.233
J	0.2	1.000	1.000	1.000
	0.6	0.996	0.991	0.997
	0.8	0.984	0.972	0.989
	1.0	0.947	0.933	0.968
	1.2	0.871	0.852	0.916
	2.0	0.478	0.384	0.498
	3.0	0.237	0.173	0.238
K	0.2	1.000	1.000	1.000
	0.6	0.997	0.994	0.998
	0.8	0.985	0.982	0.991
	1.0	0.956	0.956	0.972
	1.2	0.894	0.898	0.922
	2.0	0.459	0.377	0.468
	3.0	0.196	0.146	0.197
L	0.2	1.000	1.000	1.000
	0.6	1.000	0.999	1.000
	0.8	0.996	0.994	0.998
	1.0	0.983	0.981	0.993
	1.2	0.945	0.938	0.970
	2.0	0.440	0.294	0.641
	3.0	0.133	0.081	0.133

*16 lots required to meet minimum sample assignment of Table VIII for AQL = 0.4, 1.0, 2.5. 17 lots required for AQL = 1.5.

TABLE 4. *Long range operating characteristics of three sampling schemes employing normal, tightened and reduced inspection (MIL–STD–105D general inspection level II)* – Continued

AQL = 1.5 Percent

Lot size code letter	Incoming percent defective	Fractions of lots accepted		
		MIL–STD–105D	Japanese Standards Assn.	Authors' modification
D*	0.3	0.989	0.989	0.990
	0.9	0.942	0.948	0.952
	1.2	0.900	0.911	0.917
	1.5	0.847	0.860	0.871
	1.8	0.788	0.797	0.815
	3.0	0.577	0.549	0.595
	4.5	0.419	0.367	0.424
G	0.3	0.998	0.999	0.998
	0.9	0.970	0.985	0.975
	1.2	0.933	0.955	0.943
	1.5	0.871	0.893	0.888
	1.8	0.786	0.793	0.806
	3.0	0.460	0.391	0.466
	4.5	0.255	0.197	0.255
H	0.3	1.000	1.000	1.000
	0.9	0.997	0.994	0.998
	1.2	0.985	0.978	0.992
	1.5	0.955	0.949	0.977
	1.8	0.895	0.888	0.940
	3.0	0.527	0.437	0.556
	4.5	0.268	0.197	0.269
J	0.3	1.000	1.000	1.000
	0.9	0.998	0.995	0.998
	1.2	0.988	0.984	0.993
	1.5	0.964	0.963	0.977
	1.8	0.912	0.917	0.938
	3.0	0.495	0.415	0.506
	4.5	0.216	0.161	0.217
K	0.3	1.000	1.000	1.000
	0.9	1.000	0.999	1.000
	1.2	0.997	0.996	0.999
	1.5	0.988	0.987	0.995
	1.8	0.963	0.961	0.982
	3.0	0.514	0.361	0.539
	4.5	0.166	0.101	0.166

*16 lots required to meet minimum sample assignment of Table VIII for AQL = 0.4, 1.0, 2.5. 17 lots required for AQL = 1.5.

TABLE 4. *Long range operating characteristics of three sampling schemes employing normal, tightened and reduced inspection (MIL–STD–105D general inspection level II)* — Continued

AQL = 2.5 percent

Lot size code letter	Incoming percent defective	Fractions of lots accepted		
		MIL–STD–105D	Japanese Standards Assn.	Authors' modification
C*	0.5	0.987	0.986	0.988
	1.5	0.938	0.944	0.946
	2.0	0.893	0.905	0.908
	2.5	0.838	0.850	0.858
	3.0	0.777	0.785	0.799
	5.0	0.564	0.534	0.578
	7.5	0.407	0.357	0.411
F	0.5	0.998	0.999	0.998
	1.5	0.970	0.983	0.972
	2.0	0.930	0.947	0.935
	2.5	0.859	0.871	0.870
	3.0	0.764	0.755	0.778
	5.0	0.430	0.355	0.434
	7.5	0.233	0.178	0.233
G	0.5	1.000	1.000	1.000
	1.5	0.996	0.991	0.997
	2.0	0.984	0.972	0.989
	2.5	0.946	0.933	0.967
	3.0	0.871	0.852	0.915
	5.0	0.478	0.384	0.497
	7.5	0.237	0.173	0.238
H	0.5	1.000	1.000	1.000
	1.5	0.997	0.994	0.998
	2.0	0.985	0.982	0.991
	2.5	0.956	0.956	0.972
	3.0	0.894	0.898	0.922
	5.0	0.459	0.377	0.467
	7.5	0.196	0.146	0.197
J	0.5	1.000	1.000	1.000
	1.5	1.000	0.999	1.000
	2.0	0.996	0.994	0.998
	2.5	0.983	0.981	0.992
	3.0	0.945	0.938	0.970
	5.0	0.440	0.294	0.454
	7.5	0.133	0.081	0.134

*16 lots required to meet minimum sample assignment of Table VIII for AQL = 0.4, 1.0, 2.5. 17 lots required for AQL = 1.5.

REFERENCES

[1] Brown, G. and H. Rutemiller, "A Cost Analysis of Sampling Inspection Under Military Standard 105D," Nav. Res. Log. Quart. *20*, 181 (Mar. 1973).

[2] Duncan, A. J., *Quality Control and Industrial Statistics* (R. C. Irwin and Co., Homewood, Illinois, 1965).

[3] Grant, E. L. and R. S. Leavenworth, *Statistical Quality Control* (McGraw-Hill, New York, 1972).

[4] *MIL–STD–105D 29 April, 1963 Military Standard—Sampling Procedures and Tables for Inspection by Attributes.* Washington, D.C.: United States Government Printing Office (1962).

[5] Ohmae, U., and R. Suga, "Basic Policy and Scheme on Modified MIL–STD–105D," International Conference on Quality Control, Tokyo (1969).

CALCULATION OF THE COST OF WARRANTY POLICIES AS A FUNCTION OF ESTIMATED LIFE DISTRIBUTIONS *

Wallace R. Blischke
University of Southern California
and
Ernest M. Scheuer
California State University, Northridge

ABSTRACT

Two types of warranties are analyzed. These are the free-replacement warranty, under which failed items are replaced free of charge until a specified total operating time has been achieved, and the pro rata warranty, under which items that fail prior to a specified time are replaced at pro rata cost to the buyer. Both the buyer's and seller's points of view are considered. The basis of the analysis is a comparison of warrantied and unwarrantied (otherwise identical) items with regard to long-run cost to the buyer and long-run profit to the seller.

Application of the results requires knowledge of certain characteristics of the life distribution of the items in question. Parametric and nonparametric methods of estimation of these characteristics from incomplete data are discussed. Single and multiple failure-mode situations are considered.

Some solutions to the problem are illustrated using incomplete data on failure times of an aircraft component.

1. INTRODUCTION

As individual consumers we frequently purchase goods or services which include a warranty of one kind or another. Most of these purchases are relatively small and most of us rarely pay much attention to the exact nature of the warranty nor do we think about its actual cost to us. From the point of view of the manufacturer of the item or supplier of the service, however, such considerations are very important ones since they can rather drastically affect profitability. In situations involving large, complex, costly pieces of equipment and dealings with other organizations (corporations, government agencies, and so forth) rather than individual consumers, the nature and cost of the warranty can become very important, indeed, both to the supplier and to the consumer. In this paper we shall analyze some common warranty policies from both points of view.

In analyzing warranty policies, the basic approach taken is a comparison of costs to the consumer and profits of the supplier, of warrantied versus unwarrantied items. From the point of view of the supplier, this comparison leads to the derivation of a differential price structure which will equate long-run profit in the two situations. From the point of view of the consumer, it involves a differential price structure such that, in the long run, the consumer is indifferent with regard to whether he purchases a warrantied or an unwarrantied item.

*This research was supported by the Office of Naval Research Under Contract N000014–67–A–0269–0028 Task. NR042–323 Code 436. Reproduction in whole or in part is permitted for any purpose of the United States Government.

The key random variable in the analysis is, of course, the number, N, of replacements required. This is a function of the type and length of the warranty, and, most importantly, of the relationship of the latter to the life distribution of the item. Thus a statistical approach to the problem begins with an analysis of life-test data. With regard to this aspect of the analysis, we shall be concerned mainly with the case of incomplete data, that is, data which consist of lifetimes of failed items and service times of an additional set of items which had not yet failed at the time at which the data were collected. Data of this type are often encountered in analyzing large, complex pieces of equipment. In such cases only a relatively few items are manufactured (say hundreds or a few thousand rather than hundreds of thousands) and it is obviously not feasible to put large numbers of items on test and observe their failure times. The only adequate set of data available are therefore field data and it is important to include in the analysis the information available from unfailed as well as failed items.

We shall look at estimation of the characteristics of the distribution of lifetimes given incomplete data and the use of these estimates in characterizing the distribution of N and in comparing costs. Nonparametric and parametric procedures will be considered. The analysis will then be extended to the case of multiple failure modes under the assumption that the several modes are independent. An example will be given in which the techniques developed are applied to data on a major aircraft component.

2. ANALYSIS OF WARRANTY POLICIES

In this section we describe and analyze two warranty policies, the free-replacement warranty and the pro rata warranty, in common use for stochastically failing items.

In the free replacement warranty the supplier will, for a one-time cost of C_1, furnish the customer with as many items as necessary to yield service for the warranted service duration, W_1.

The pro rata warranty operates as follows. Upon failure of an item at age X, a new item is furnished at a cost

$$C_2[\min(W_2, X)/W_2]. \tag{2.1}$$

The replacement item is warranted anew. That is, if the lifetime of the item at its failure exceeds the length W_2 of the warranty period, a new item must be purchased at the full price of C_2; while if the life time of the item at its failure is less than the length of the warranty period, a new item is sold at a fraction of the full price, the fraction being the percentage of the warranty period experienced by the old item.

We will compare these two warranty policies with each other and with the situation prevailing if no warranty of any kind is in effect and a new item must be bought at a cost of K, no matter what the lifetime of its predecessor was. The interests of the customer will be considered first and those of the supplier will be considered separately.

The following notation will be used:

X = lifetime of the item (a nonnegative random variable);
L = life cycle of the item;
$F(\cdot)$ = cumulative distribution function of X;
$\bar{F}(t) = 1 - F(t)$;
$X_1, X_2 \ldots$ = lifetimes of successive items (independently and identically distributed as F);
$S_n = \sum_{i=1}^{n} X_i$; $S_0 \equiv 0$;

$N = N(t)$ = number of replacements of the item made in the time interval $[0, t]$
$= \max\{n : S_n \leq t\}$;
$M = M(t) = E[N(t)]$ (The "renewal function");
$\mu = E(X)$;
$\mu_W = \int_0^W x\, dF(x)$ (The "partial expectation of X");
$\gamma(t)$ = remaining lifetime of the item in operation at time t
$= S_{N(t)+1} - t$;

and

$F^{(n)}(\cdot) = n$-fold convolution of $F(\cdot)$ with itself.

We know from renewal theory that

$$P[N(t) = n] = F^{(n)}(t) - F^{(n+1)}(t), \qquad n = 0, 1, 2, \ldots$$

where

$$F^{(0)}(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}.$$

and

$$M(t) = \sum_{n=1}^{\infty} F^{(n)}(t).$$

We now proceed to give an economic analysis of the two warranty policies.

2.1 The Customer's Point of View

2.1.1 The Free-Replacement Warranty

An important observation to make concerning the free replacement warranty is that when the warranty expires at time W_1, the customer does not immediately incur the replacement cost, C_1. That cost is not incurred until the failure of the item then in service, at time (say) $Y > W_1$, where $Y = W_1 + \gamma(W_1)$.

It is easy to prove (see, e.g., Smith [7], Equation (1.1)) that

$$E(Y) = \mu[1 + M(W_1)]. \tag{2.2}$$

Thus, over the life cycle of the item the average number of times payment must be made under the free-replacement warranty is $1 + L/E(Y) = 1 + L/\{\mu[1 + M(W_1)]\}$ † and therefore the average total life cycle cost of the item with a free-replacement warranty is

† A sharper result is that the expected number of times payment must be made under the free-replacement warranty is $1 + M_Y(L)$, where $M_Y(\cdot)$ is the renewal function for the random variable Y. Although an expression for $M_Y(\cdot)$ is not known, we know from the Elementary Renewal Theorem that $M_Y(t)/t \to 1/E(Y)$. The reader is cautioned, however, that the result is an asymptotic one and may be valid only for relatively large L.

$$C_1(1+L/\{\mu[1+M(W_1)]\}). \tag{2.3}$$

If no warranty is in force, a cost of K is incurred, on the average, $1+M(L)$ times, so that over the life cycle the average total cost is $K[1+M(L)]$. Equating the latter to (2.3) we obtain for the cost under free-replacement warranty at which the customer would be indifferent between buying the item with or without a warranty the expression

$$C_1^*=\frac{\mu K[1+M(L)][1+M(W_1)]}{\mu[1+M(W_1)]+L}. \tag{2.4}$$

If $C_1>(<)C_1^*$, the customer is better (worse) off to buy the item without (with) the free-replacement warranty than with (without) it.

Note, incidentally, that if L is large relative to μ, $M(L)$ is approximately L/μ, in which case

$$C_1^* \approx \frac{\mu K\left(1+\frac{L}{\mu}\right)[1+M(W_1)]}{\mu[1+M(W_1)]+L}.$$

As $L\to\infty$, this becomes $C_1^*=K[1+M(W_1)]$.

Another important point in this connection is that if the length of the item's life cycle is sufficiently long, it may be worthwhile to include in the preceding analysis discounting of future payments at rate α (say) back to time zero. With the free-replacement warranty, payments of C_1 are made, on the average, at time intervals of length $E(Y)$. Thus the average present value of all payments into the indefinite future (i.e., $L\to\infty$) *is*

$$\begin{aligned} C_1\sum_{j=0}^{\infty}\exp\,[-j\alpha E(Y)] &= C_1\{1-\exp\,[-\alpha E(Y)]\}^{-1} \\ &= C_1\{1-\exp\,(-\alpha\mu[1+M(W_1)]\}^{-1}. \end{aligned} \tag{2.5}$$

Similarly, if the payments of K that the customer makes, on average, at intervals of duration μ when no warranty is in effect are discounted, the average present value of all these payments into the indefinite future is

$$K\sum_{j=0}^{\infty}\exp\,(-j\alpha\mu)=K[1-\exp\,(-\alpha\mu)]^{-1}. \tag{2.6}$$

Equating (2.5) and (2.6) yields

$$C_1^{**}=\frac{K\{1-\exp\,(-\alpha\mu[1+M(W_1)])\}}{1-\exp\,(-\alpha\mu)} \tag{2.7}$$

as the price at which the buyer would be indifferent between buying the item with or without a free-

replacement warranty when the present value of future payments is considered.

If $C_1 > (<)$**, the customer is better (worse) off to buy the item without (with) the free-replacement warranty that with (without) it when present value considerations enter.

2.1.2 The Pro Rata Warranty

We first need to obtain the expected cost, say R, of the $(i+1)$st replacement item. This is given by

$$R \equiv \frac{C_2}{W_2} E[\min(W_2, X_i)] = \frac{C_2}{W_2}\left[\int_0^{W_2} x dF(x) + \int_{W_2}^{\infty} W_2 dF(x)\right] \tag{2.8}$$

$$= \frac{C_2}{W_2}[\mu_{W_2} + W_2 \bar{F}(W_2)].$$

Conditional upon there being exactly n replacements over the life cycle of the item (i.e., upon $N(L) = n$), the expected cost over the life cycle is

$$C_2(1+nR) = C_2\left\{1 + \frac{n}{W_2}[\mu_{W_2} + W_2\bar{F}(W_2)]\right\}.$$

Unconditioning on $N(L)$, we obtain for the expected cost over the life cycle under pro rata warranty

$$C_2\left\{1 + \frac{M(L)}{W_2}[\mu_{W_2} + W_2\bar{F}(W_2)]\right\}. \tag{2.9}$$

Without warranty, the expected cost over the life cycle is $K[1+M(L)]$. Equating this to (2.9), we obtain for the cost C_2^* at which the customer would be indifferent between buying the item with or without pro rata warranty the expression

$$C_2^* = \frac{W_2 K[1+M(L)]}{W_2 + M(L)[\mu_{W_2} + W_2\bar{F}(W_2)]}. \tag{2.10}$$

If $C_2 > (<)\ C_2^*$ the customer is better (worse) off to buy the item without (with) the pro rata warranty than with (without) it.

If, instead of adding up all the costs, we take the sum of their present values (with discount rate α) back to the beginning of the life cycle, we obtain for the average present value of all payments into the indefinite future (i.e., $L \to \infty$) under pro rata warranty

$$R\sum_{j=0}^{\infty} \exp[-j\alpha\mu] = R[1-\exp(-\alpha\mu)]^{-1} \tag{2.11}$$

$$= \frac{C_2}{W_2}[\mu_{W_2} + W_2\bar{F}(W_2)][1-\exp(-\alpha\mu)]^{-1}.$$

The average present value of all future payments in the no-warranty situation is given by (2.6). Equating (2.6) and (2.11) gives

$$C_2^* = \frac{KW_2}{\mu_{W_2} + W_2\bar{F}(W_2)} \tag{2.12}$$

as the price at which the buyer would be indifferent between buying the item with or without a pro rata warranty when the present value of future payments is considered.

If $C_2 > (<) \; C_2^{**}$ the customer is better (worse) off buying the item without (with) the pro rata warranty than with (without) it when present value considerations enter.

2.1.3 Comparison of the Free-Replacement and Pro Rata Warranty

If the customer has a choice between the free-replacement and pro rata warranty, he can make that choice by comparing the average total life cycle costs, given by (2.3) and (2.9), respectively, or the average total discounted costs, given by (2.5) and (2.11), respectively.

Note that C_2^{**} is not a function of the discount rate, α. This is so because payments are made at the same time with or without the pro rata warranty. In fact, it is easily seen that as $L \to \infty$, C_2^* (which does not involve discounting) approaches C_2^{**} (which does involve discounting).

2.2 The Seller's Point of View

The foregoing discussion has focused on the customer and has concerned itself solely with the costs he will bear under the free-replacement, pro rata, and no-warranty situations.

We turn now to consider the seller's point of view, specifically his profit. This profit is simply the difference between his revenue (the cost to the customer) and his costs. In the free-replacement warranty the seller's revenue per warranty period is fixed (at C_1) while his costs are a random variable since the number of free replacements he must make are random. We will obtain the expected value of these costs. In the pro rata warranty it is the revenue per transaction that is a random variable. Its expected value, R, is given in (2.8).

In the free-replacement warranty the length of time between payments is the random variable Y. This can be expressed as

$$Y = W_1 + \gamma(W_1) = S_{N(W_1)+1}.$$

From this expression we see that the expected number of units supplied over the period Y is $M(W_1) + 1$. (This includes the initial one that was purchased and the subsequent free replacements.)

If we let g denote the seller's cost per item and $T_1(T_2)$ the expected profit per warranty cycle (transaction) under the free-replacement (pro rata) warranty, we see from the preceding discussion that

$$T_1 = C_1 - g[M(W_1) + 1] \tag{2.13}$$

and

$$T_2 = R - g = \frac{C_2}{W_2}[\mu_{W_2} + W_2\bar{F}(W_2)] - g. \tag{2.14}$$

Now we can obtain average total profit or the sum of present value of average profits under the free-replacement or pro rata warranties by replacing C_1 and C_2 by T_1 and T_2 in Equations (2.3), (2.5), (2.9), and (2.11) as appropriate. Similarly, we obtain the selling prices for free-replacement and pro rata warranties at which the seller is indifferent between selling with or without warranty using equations analogous to (2.4), (2.7), (2.10), or (2.12).

3. ESTIMATION OF WARRANTY COSTS ASSUMING A SINGLE FAILURE MODE

The preceding analysis assumed knowledge of the time-to-failure cumulative distribution function $F(\cdot)$ of the item in question, as well as its mean μ, and the probability distribution of the associated renewal process, $N(\cdot)$, which in turn requires the successive convolutions, $F^{(k)}(\cdot)$, of $F(\cdot)$ with itself. Finally, to complete the cost analysis of the two warranties, we require, in addition, μ_{W_2} and the renewal function $M(\cdot)=E[N(\cdot)]$. In principle, knowledge of $F(\cdot)$ is sufficient to know all the other quantities.

In most practical situations $F(\cdot)$ will not be known. In some situations the form of $F(\cdot)$ may be known (e.g., normal, exponential, Weibull, or the like) but not the specific parameters. In other situations not even the form of $F(\cdot)$ is known. In the former situation the parameters must be estimated; in the latter, one may either follow a nonparametric approach or first try to elicit the form of $F(\cdot)$ and then estimate the parameters within that form. How these estimation tasks are performed depends upon the kind of data available.

In this section the estimation procedures, assuming a single failure mode, will be given; these results will be extended to multiple failure modes in the next section. In all of the remainder of this section we shall continue to stress general results, but in the parametric case will use the normal distribution to illustrate the techniques. This choice was made on the basis of the apparent normality of the data in the application which follows. We will describe some procedures that can be used quite generally with "incomplete" data as well as with "complete" samples.

A set of incomplete lifetime data consists of the time-to-failure of some items, periods of observation on some items which may or may not begin at the inception of use of the item and may or may not end with the failure of the item. A special case of this rather general sampling situation is the complete sample in which some number of items is observed from their first use until their failure and the time-to-failure of each is noted.

3.1 Nonparametric Estimation

Kaplan and Meier [3] provide methods for the nonparametric estimation of $F(\cdot)$ from incomplete data. Their estimate $\hat{F}(\cdot)$ of $F(\cdot)$ is a step function; for complete samples it coincides with the usual empirical cumulative distribution function. With the estimate of $F(\cdot)$, one can in turn estimate $\mu=E^{\circ}X'$, $\sigma^2=\text{var}\,(X)$ (if needed), $E[\min(X, W)]$, $F^{(n)}(\cdot)$, $P[N(t)=n]$, $M(t)$, etc., from the formulas, valid for nonnegative random variables,

$$\mu=\int_0^{\infty}[1-F(t)]dt, \tag{3.1}$$

$$\hat{\sigma}^2=2\int_0^{\infty}t[1-F(t)]dt-\mu^2, \tag{3.2}$$

$$E[\min(X, W)] = \int_0^W [1-F(t)]dt, \tag{3.3}$$

$$F^{(n)}(t) = \int_0^t F^{(n-1)}(t-x)dF(x), \qquad n=2, 3, \ldots \tag{3.4}$$

$$P[N(t)=n] = F^{(n)}(t) - F^{(n+1)}(t), \qquad n=0, 1, 2, \ldots \tag{3.5}$$

$$M(t) = E[N(t)] = \sum_{n=0}^{\infty} nP[N(t)=n] = \sum_{n=1}^{\infty} F^{(n)}(t). \tag{3.6}$$

Substituting the step function $\hat{F}(\cdot)$ in these formulas yields sums in place of integrals in the equations for $\hat{\mu}$, $\hat{\sigma}^2$, $\hat{E}[\min(X, W)]$, and $\hat{F}^{(n)}(t)$.

3.2 Estimation Under Parametric Assumptions

If one wishes to go beyond a nonparametric approach to the problem without *assuming* a distributional form for the time-to-failure distribution, he must elicit that form somehow and then estimate parameters. A popular method for doing these two tasks at essentially the same time is probability plotting. With this technique the data are plotted on special graph papers which are specific to certain distributional forms. A nearly linear plot on a particular kind of probability paper is evidence that the data came from the associated distributional family. Once having chosen a particular distribution, the graph can also be used to estimate parameters. This is done by reading off the slope of the best-fit line, its intercept on some axis or an associated scale, or the like. Details can be found, e.g., in Nelson [5, 6]. In these papers the normal, lognormal, exponential, Weibull, and gamma distributions are discussed.

For certain distributions maximum likelihood estimates are also available. For the normal distribution, the results are given by Cohen [1, 2]. In general, the likelihood function for a sample of size n_T in which there are n_1 failures, at times $X_{11}, \ldots, X_{1n_1}$, say, and $n_2 = n_T - n_1$ unfailed items, with service times $X_{21}, \ldots, X_{2n_2}$, is

$$L = \left(\prod_{i=1}^{n_1} f(X_{1i}) \right) \left(\prod_{i=1}^{n_2} \bar{F}(X_{2i}) \right), \tag{3.7}$$

where $F(\cdot)$ is the assumed *CDF* and $f(\cdot)$ the corresponding density. For most distributions the resulting likelihood equations are quite complex and numerical methods are required for their solution.

For the normal distribution the likelihood equations are

$$0 = \frac{\partial \log L}{\partial \mu} = \sum_{i=1}^{n_1} \left(\frac{X_{1i}-\mu}{\sigma^2} \right) + \sum_{i=1}^{n_2} \frac{f(x_{2i})}{\bar{F}(X_{2i})} \tag{3.8}$$

$$0=\frac{\partial \log L}{\partial \sigma}=\frac{1}{\sigma^2}\sum_{i=1}^{n_1}(X_{1i}-\mu)^2-\frac{n_1}{\sigma}+\frac{1}{\sigma}\sum_{i=1}^{n_2}\frac{(X_{2i}-\mu)f(X_{2i})}{\bar{F}(X_{2i})} \tag{3.9}$$

Although an analytical solution is not possible, these equations can be solved rather easily by computer methods.

Incidentally, pursuant to the usual sequence of events in maximum likelihood estimation, one can next determine the second partials of log L. For the normal distribution, these are

$$\frac{\partial^2 \log L}{\partial \mu^2}=\frac{-n_1}{\sigma^2}+\frac{1}{\sigma^2}\sum_{i=1}^{n_2}\frac{(X_{2i}-\mu)f(X_{2i})}{\bar{F}(X_{2i})}-\sum_{i=1}^{n_2}\frac{f^2(X_{2i})}{\bar{F}^2(X_{2i})}, \tag{3.10}$$

$$\frac{\partial^2 \log L}{\partial \mu\,\partial \sigma}=\frac{-2}{\sigma^3}\sum_{i=1}^{n_1}(X_{1i}-\mu)+\frac{1}{\sigma^3}\sum_{i=1}^{n_1}\frac{[(X_{2i}-\mu)^2-\sigma^2]f(X_{2i})}{\bar{F}(X_{2i})}-\frac{1}{\sigma}\sum_{i=1}^{n_2}\frac{(X_{2i}-\mu)f^2(X_{2i})}{\bar{F}^2(X_{2i})}, \tag{3.11}$$

and

$$\frac{\partial^2 \log L}{\partial \sigma^2}=\frac{-3}{\sigma^4}\sum_{i=1}^{n_1}(X_{1i}-\mu)^2+\frac{n_1}{\sigma^2}-\frac{2}{\sigma^2}\sum_{i=1}^{n_2}\frac{(X_{2i}-\mu)f(X_{2i})}{\bar{F}(X_{2i})} + \frac{1}{\sigma^4}\sum_{i=1}^{n_2}\frac{(X_{2i}-\mu)^3 f(X_{2i})}{\bar{F}(X_{2i})}-\frac{1}{\sigma^2}\sum_{i=1}^{n_2}\frac{(X_{2i}-\mu)^2 f^2(X_{2i})}{\bar{F}^2(X_{2i})}. \tag{3.12}$$

As usual, these results can be used to estimate the asymptotic variance of the estimates. Conceptually this could be done either by determining the expected values of the derivatives or by evaluating them at the observed data points. Because of the nature of the likelihood function however, some difficulties are encountered in attempting to determine the required expectations when the data are incomplete. As usual, the expectations would be taken conditional upon the (random) number of failed and unfailed items in the sample. The density of service times of unfailed items, however, apparently depends on the times at which items were put into service, which could vary greatly in different applications. In the application to be discussed below we therefore use the second approach, as did Cohen [2] in his example.

Once μ and σ are estimated, it is easy to estimate $F^{(n)}$, and thus the probability distribution of N_1 in the normal case. From this, estimates of M and the variance of N are easily calculated. Conceptually, one could proceed to use a normal approximation to obtain an approximate confidence interval for M based on estimates of the information matrix and the standard results concerning asymptotic variances of functions of estimators. The estimator of M is a very complicated function of the estimators of μ and σ, however, and we have not pursued the analysis to this point; we plan to do so in future research. In any case, the estimate of M enables one to calculate at least a point estimate of the cost associated with Warranty 2.

To obtain a similar result for Warranty 1 an estimate of $E(\min(X, W))$ is required. In the normal case, this can be obtained easily by using an alternate form of the expression in (3.3), namely

$$E(\min(X, W))=\int_{-\infty}^{W} xf(x)\,dx+W\bar{F}(W)$$

$$= \int_{-\infty}^{W} \frac{x}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} dx + W\bar{F}(W)$$

$$= -\sigma^2 f(W) + \mu F(W) + W\bar{F}(W).$$

4. ANALYSIS WITH MULTIPLE INDEPENDENT FAILURE MODES

An important extension of the above analysis is to the case in which there are multiple competing modes of failure. If the causes of failure act independently of one another, the extension of the analysis is conceptually straightforward. If they do not, additional complications are encountered. (With regard to the estimation problem, see for example, Moeschberger [4] and the references he cites. The analysis of the random variable N is even more complex.) Here we shall be concerned with the case of independent failure modes, and will ultimately assume that each leads to a normal distribution of lifetimes. It will be shown that under the assumption of independence the maximum likelihood procedure in the multimode case involves the same basic computations as in the single mode case. It should be noted that the same is true of the graphical and nonparametric procedures.

4.1 Parameter Estimation

Suppose first that there are two independent failure modes with *CDF*'s of times to failure, F_1 and F_2, and corresponding densities, f_1 and f_2. Suppose further, that in a sample of size n_T, n_1 items are observed to fail by mode 1, n_2 by node 2 and $n_3 = n_T - n_1 - n_2$ items do not fail. Let $X_{11}, \ldots, X_{1n_1}$ be the lifetimes of the items that fail by mode 1, $X_{21}, \ldots, X_{2n_2}$ the lifetimes of the items that fail by mode 2 and $X_{31}, \ldots, X_{3n_3}$ the service times of the unfailed items. Then, given n_1, n_2 and n_3, the likelihood function is

$$L = \left(\prod_{j=1}^{n_1} f_1(x_{1j})\bar{F}_2(X_{1j})\right)\left(\prod_{j=1}^{n_2} f_2(x_{2j})\bar{F}_1(x_{2j})\right)\left(\prod_{j=1}^{n_3} \bar{F}_1(x_{3j})\bar{F}_2(X_{3j})\right), \tag{4.1}$$

where any product is taken to be unity if the corresponding n_i is 0.

The likelihood equations and elements of the information matrix are again expressed in terms of the derivatives of the $f_i(\cdot)$ and $F_i(\cdot)$ with respect to the parameters of the distributions. Assume now the normal case with $f_i(\cdot)$, $F_i(\cdot)$ having parameters μ_i, σ_i^2. For simplicity, denote the required derivatives $f_{i,\mu}(\cdot)$, $F_{i,\mu}(\cdot)$, $F_{i,\mu\mu}(\cdot)$, and so forth. (All of these appear in Equations (3.8) through (3.12).) The likelihood equations are

$$0 = \frac{\partial \log L}{\partial \mu_1} = \sum_{j=1}^{n_1} \frac{f_{1,\mu}(X_{1j})}{f_1(X_{1j})} - \sum_{j=1}^{n_2} \frac{F_{1,\mu}(X_{2j})}{\bar{F}_1(X_{2j})} - \sum_{j=1}^{n_3} \frac{F_{1,\mu}(X_{3j})}{\bar{F}_1(X_{3j})}$$

$$0 = \frac{\partial \log L}{\partial \sigma_1} = \sum_{j=1}^{n_1} \frac{f_{1,\sigma}(X_{1j})}{f_1(X_{1j})} - \sum_{j=1}^{n_2} \frac{F_{1,\sigma}(X_{2j})}{\bar{F}_1(X_{2j})} - \sum_{j=1}^{n_3} \frac{F_{1,\sigma}(X_{3j})}{\bar{F}_1(X_{3j})}$$

$$0 = \frac{\partial \log L}{\partial \mu_2} = \sum_{j=1}^{n_1} \frac{F_{2,\mu}(X_{1j})}{\bar{F}_2(X_{1j})} + \sum_{j=1}^{n_2} \frac{f_{2,\mu}(X_{2j})}{f_2(X_{2j})} - \sum_{j=1}^{n_3} \frac{F_{2,\mu}(X_{3j})}{\bar{F}_2(X_{3j})}$$

4.2) $$0=\frac{\partial \log L}{\partial \sigma_2}=\sum_{j=1}^{n_1}\frac{F_{2,\sigma}(X_{1j})}{\bar{F}_2(X_{1j})}+\sum_{j=1}^{n_2}\frac{f_{2,\sigma}(X_{2j})}{f_2(X_{2j})}-\sum_{j=1}^{n_3}\frac{F_{2,\sigma}(X_{3j})}{\bar{F}_2(X_{3j})},$$

where $\bar{F}.(\cdot)=1-F.(\cdot)$ as before. Note that the likelihood equations are separable; the first pair involve only $f_1(\cdot)$, $F_1(\cdot)$ and their derivatives, the second pair only $f_2(\cdot)$, $F_2(\cdot)$ and their derivatives. This will extend to any number of failure modes. It is easily seen that under independence the likelihood equations will always partition into pairs, the ith pair of which will involve only μ_i and σ_i. Furthermore, in the ith pair only those items which failed by mode i will be treated as failed items; all other observations will be treated as service times. (A similar result, of course, holds for distributions other than the normal as well and also for combinations of distributions—e.g., one mode leading to normally distributed lifetimes and another to exponential—as long as the causes act independently). It follows that if one can analyze data assuming a single mode of failure and a particular distribution, multiple independent modes can be handled just as easily.

Another important conclusion is that the information matrix can be handled just as easily for the multiple-mode as for the single-mode situation when causes are independent. This is easily seen by observing that the mixed second derivatives of the logarithm of the likelihood will be zero whenever parameters from two different distributions are involved. It follows that the information matrix will have submatrices with entries corresponding to the individual distributions on the main diagonal (in the normal case these will be 2×2 matrices) and zeroes elsewhere, so that if the information matrix can be evaluated and inverted in the case of a single mode, this can also be done in the case of several. Thus, under independence, the existence of multiple failure modes introduces no additional complexity as far as estimation of the parameters of the life distribution is concerned.

4.2 Distribution of $N(t)$ in the Case of Multiple Modes

The distribution of $N(t)$ discussed in the second section of this paper was derived under the assumption of a single mode of failure. The extension of the analysis to the multiple-mode case is again conceptually easy if the causes are independent. The basic result, $P(N(t)=n)=F^{(n)}(t)-F^{(n+1)}(t)$, still holds in the multiple-mode case, but $F^{(n)}$ takes on a rather different form. Consider the case of r failure modes. The form of $F^{(n)}$ is most easily seen by noting that, given independent causes, the situation is exactly analogous to that in which a system fails, and is replaced in its entirety, when the first of r independently operating components fails. In the analysis of $N(t)$, $F(\cdot)$ now becomes the CDF of the minimum of r independent, but nonidentically distributed random variables and $F^{(n)}(\cdot)$ the n-fold convolution of such CDF's. Note, incidentally, that the relevant distribution here is the unconditional distribution of the random variable in question whereas the likelihood function of Equation 4.1) was conditional upon n_1, n_2, and n_3. This accounts for some apparent dissimilarities between the results to follow and those of the previous section. (A similar situation, of course, prevailed in the single-mode case.)

From a theoretical point of view, the distributions required are well-known results. Let $X_1, \ldots, X_r$ be the random lifetimes corresponding to the r failure modes, $F_1(\cdot), \ldots, F_r(\cdot)$ the corresponding CDF's and let $Z=\min(X_1, \ldots, X_r)$. The CDF of Z is then

4.3) $$F_Z(z)=1-\prod_{i=1}^{r}[1-F_i(z)],$$

with corresponding density

$$f_Z(z) = \left(\sum_{j=1}^{r} \frac{f_j(z)}{\bar{F}_j(z)}\right) \bar{F}_Z(z). \tag{4.4}$$

Furthermore, under independence the convolution $F^{(n)}(t)$ is given by the iterated integral

$$F^{(n)}(t) = \int_0^t \int_0^{t-z_1} \cdots \int_0^{t-z_1-\cdots-z_{k-1}} \prod_{i=1}^{k} [f_Z(z_i)dz_i], \tag{4.5}$$

which, of course, is ordinarily expressed and calculated recursively.

Unfortunately, in implementation this theoretical result leads to explicit expressions for the convolution only in a few relatively simple cases, the exponential distribution being perhaps the only such case of interest in dealing with life-test data. For the normal distribution, the iterated integral of Equation (4.5) can only be evaluated numerically and then with some difficulty, particularly as r increases. In fact, the reason that a numerical approach is feasible at all is that in nearly all applications one would expect N to be small with high probability. (It seems unlikely that a warranty that would require many replacements very often would be established.)

4.3 Estimation of $E(\min(W_2, Z))$

Since Z is the random lifetime of an item in the multimode case, the characteristic of interest in analyzing the pro rata warranty is $E(\min(W_2, Z)) = E(\min(W_2, X_1, \ldots, X_r))$. This is obtained by substituting $F_Z(\cdot)$ for $F(\cdot)$ in the integrals of Equation (2.8). In the normal case numerical integration is required to calculate the values of μ_{W_2} and $\bar{F}(W_2)$. The nonparametric approach can also be used here by making the analogous substitution with sample *CDF*'s.

Incidentally, another expectation of interest is $E(Z)$ itself. For $r=2$, this quantity is

$$E(Z) = \mu_1 + \mu_2 - \int_{-\infty}^{\infty} z[f_1(z)F_2(z) + f_2(z)F_1(z)]dz. \tag{4.6}$$

The general expression involves $\sum_{i=1}^{r} \mu_i$ and all integrals of the form $\int z f_i(z) F_{j_1}(z) \cdots F_{j_\nu}(z)dz$ for $\nu = 1, \ldots, r$. Incidentally, it is easily shown that $E(Z) < \min(\mu_1, \ldots, \mu_r)$.

5. NUMERICAL EXAMPLE

In a study of aircraft components, subjected to a variety of stresses in flight, two basic modes of failure were observed. (Each was actually broken into several submodes, but for simplicity we shall ignore these finer distinctions.) The incomplete data included 174 observations, of which 22 were failure times for items which failed by Mode 1, 25 were failure times for items which failed by Mode 2, and 127 were service times for unfailed items. The observations are in hundreds of hours of flight time.

The data for each of the two modes as well as the composite failure data were plotted on normal, lognormal, exponential, Weibull, and gamma hazard paper. Somewhat unexpectedly, the normal plot consistently yielded the best fit.

5.1 Estimates of Parameters

For purposes of comparison, the graphical and maximum likelihood estimates of μ and σ in the normal case and the nonparametric estimates of μ and σ were calculated for each of the two failure modes as well as for the composite. The results are given in Table 1. (The graphical estimates are given to only the nearest integer because it is unlikely that they can be read more accurately than that.)

TABLE 1. *Normal and nonparametric estimates of* μ and σ

Failure mode	Normal				Nonparametric	
	Graphical		Maximum likelihood			
	μ	σ	μ	σ	μ	σ
1	80	24	76.1	25.4	73.6	20.1
2	72	19	70.5	20.6	69.3	18.5
Both	61	20	60.4	19.3	60.7	19.1

The results are reasonably consistent. The only apparent pattern, which may, of course, be unique to this set of data, is that in every analysis (including those involving submodes not reported here) the nonparametric estimate of σ is less than either estimate under the assumption of normality. Note, incidentally, that in every case the estimate of μ is about three times the estimate of σ or greater. Thus virtually the entire distribution lies on the positive part of the axis. One would, or course, prefer that μ be at least, say, 5σ to justify normality but this result in addition to the fact that a nearly linear relationship was observed in the plots on normal hazard paper makes the normality assumption at least not patently untenable.

The program written to calculate the maximum likelihood estimates also evaluates estimates of the information matrix and its inverse under the assumption of normality. The elements of the latter, denoted I_{11}, I_{12}, and I_{22}, are given in Table 2. These can be used in the usual fashion to obtain

TABLE 2. *Estimates of the elements of the inverse of the information matrix*

Mode	I_{11}	I_{12}	I_{22}
1	30.33	16.72	14.55
2	15.55	8.72	7.37
Both	6.36	2.76	3.65

approximate confidence intervals for or tests of hypotheses about the parameters. For example, an approximate test of the hypothesis that the two failure types have the same *MTTF* yields a t-value of $(76.1-70.5)/\sqrt{30.33+15.55}=0.83$.† Since the standard deviations are also quite similar we shall complete the computations of the cost analysis under the assumption of a single mode of failure.

†It should be noted that we cannot provide a theoretical justification for using a Student-t in making this comparison. It is included merely as an indication that there are apparently no substantial differences between the two *MTTF*'s.

5.2 Cost Analysis

5.2.1 The Free Replacement Warranty

In the ensuing calculations we shall assume a normal distribution and use $\mu = 60.4$ and $\sigma = 19.3$. Corresponding values of $M(t)$ are given in Table 3 for several values of t.

TABLE 3. *Values of* M(t) *for* $\mu = 60.4$ *and* $\sigma = 19.3$

t	20	30	40	50	60	70	80
$M(t)$	0.018	0.058	0.147	0.300	0.505	0.722	0.913

t	400	600	800	1000
$M(t)$	6.17	9.52	12.80	16.12

Substitution of these results into Equation (2.4) provides a comparison of K, the cost of an unwarrantied item, and C_1^* the corresponding indifference price for a warrantied item. The ratio of these two quantities is given in Table 4 for various values of L, the life cycle, and W_1, the warranty period. (The accuracy of the results for the smaller L-values is uncertain.)

TABLE 4. *Values of* C_1^*/K *for* $W_1 = 30(10)80$ *and* $L = 400(200)1000$

W_1	L			
	400	600	800	1000
30	†	1.01	1.02	1.03
40	1.06	1.09	1.10	1.11
50	1.18	1.22	1.23	1.25
60	1.33	1.38	1.41	1.43
70	1.40	1.55	1.59	1.61
80	1.61	1.70	1.74	1.77

† Approximation not adequate.

Discounted values C_1^{**}/K are given in Table 5 for annual discount rates of 0.05, 0.10, 0.15 and 0.20 and various values of W_1. In order to perform the calculations some assumption must be made about annual service times. For purposes of this example we assumed an average annual total service time of 3,000 hours. Thus the values of α to be substituted into Equation (2.7) are 0.05/30, 0.10/30, and so forth.

TABLE 5. *Values of* C_1^{**}/K *for annual discount rates of* 0.05(0.05)0.20 *and* $W_1 = 20(10)80$

W_1	Discount rate			
	0.05	0.10	0.15	0.20
20	1.017	1.016	1.015	1.015
30	1.055	1.052	1.049	1.047
40	1.139	1.131	1.123	1.116
50	1.281	1.263	1.246	1.230
60	1.468	1.433	1.401	1.371
70	1.662	1.607	1.556	1.509
80	1.829	1.753	1.684	1.620

The seller's profit can be calculated by substituting the above results into Equation (2.13).

5.2.2 The Pro Rata Warranty

The pro rata warranty is analyzed by estimating $E[\min(W_2, X)]$. Estimates of this quantity and the resulting estimates of C_2^*/K for the above example (under the assumption of normality) are given in Table 6 for various values of W_2. The corresponding discounted values are given in Table 7. As one might expect, the ratios are less for the pro rata than for the free-replacement warranty. The same is true of discounted values under the two warranties. Again, the results from the point of view of the seller can be obtained from Equation (2.14).

TABLE 6. *Values of* C_2^*/K *for* $L = 400(200)1000$ *and* $W_2 = 20(10)80$

W_2	$E[\min(X, W_2)]$	400	600	800	1000
20	19.9	1.004	1.005	1.005	1.005
30	29.5	1.015	1.015	1.016	1.016
40	38.6	1.031	1.033	1.034	1.034
50	46.4	1.066	1.070	1.072	1.073
60	52.5	1.121	1.128	1.131	1.133
70	56.6	1.197	1.210	1.216	1.220
80	58.8	1.295	1.316	1.326	1.333

TABLE 7. *Discounted values of* C_2^{**}/K *for* $W_2 = 20(10)80$

W_2	20	30	40	50	60	70	80
C_2^{**}/K	1.005	1.017	1.036	1.078	1.143	1.237	1.361

ACKNOWLEDGMENTS

The authors wish to express their appreciation to Clark M. Roberts, who programmed most of the computations of the example, to the Office of Naval Research for its financial support, and to a referee for valuable comments.

REFERENCES

[1] Cohen, A. Clifford, Jr., "Progressively Censored Samples in Life Testing," Technometrics *5*, 327–339 (1963).

[2] Cohen, A. Clifford, "Life Testing and Early Failure," Technometrics *8*, 539–545 (1966).

[3] Kaplan, E. L. and Paul Meier, "Nonparametric Estimation From Incomplete Samples," J. Am. Statist. Assn. *53*, 457–481 (1958).

[4] Moeschberger, M. L., "Life Tests Under Dependent Competing Causes of Failure," Technometrics *16*, 39–47 (1974).

[5] Nelson, Wayne, "Hazard Plotting For Incomplete Failure Data," Journal of Quality Technology *1*, 27–52 (1969).

[6] Nelson, Wayne, "Theory and Application of Hazard Plotting for Censored Data," Technometrics, *14*, 945–966 (1972).

[7] Smith, Walter L., "Renewal Theory and Its Ramifications," J. Royal Statist. Soc. (B), *20*, 243–302 (1958), (with discussion).

CONTROL CHARTS FOR EXPONENTIALLY DISTRIBUTED PRODUCT LIFE

Tuncel M. Yegulalp

Henry Krumb School of Mines
Columbia University

ABSTRACT

This paper introduces a special control chart procedure for exponentially distributed product life. Statistical control of product life in manufacturing requires continuing life tests of manufactured product so as to detect changes in product life and take appropriate corrective action. These life testing experiments may become exceedingly time consuming and thus can be both impractical because of serious time delays in implementing corrective action on the process when indicated, and quite uneconomical.

It is desirable to inquire into the character of life testing by means of a control chart procedure based on the real time to first failure within given samples. Measuring the real minimum life provides a considerable reduction in duration of the testing procedure and in the number of specimens destroyed, yielding a considerable economy over the Shewhart's $\bar{X}$ control chart.

I. INTRODUCTION

When product life is an important characteristic of product quality, control charts on the manufacturing process are needed in order to implement effective statistical control so that the product life may be maintained at a desired level. The Shewhart control chart technique, which has been found universally applicable because if does not require knowledge of the form of the initial distribution of the random variable in question may seem to be useful in life testing. But, in order to maintain a control chart on product life one naturally needs the results of continuing life testing experiments which may become exceedingly time consuming and thus can be both impractical because of serious time delays in implementing corrective action on the process when indicated, and quite uneconomical. Hence other approaches have been sought.

Accelerated life testing has become a favorite practice among manufacturers in order to shorten the duration of life testings required and so, hopefully reduce related costs. This procedure has not provided effective process control, since the relationship between the results of accelerated life tests and the magnitude of true life is rarely accurately established. Furthermore, the procedure for accelerated life testing does not readily lend itself to the detection of assignable causes of variation. It seems desirable to inquire into the character of life testing by means of a control chart procedure based on the real time to first failure within given samples. Thus, measuring the real minimum life provides a considerable reduction in duration of the testing procedure and in the number of specimens destroyed, yielding a considerable economy over the conventional $\bar{X}$ control chart which is impractical for life testing.

One may ask: Is it possible to design a procedure where one can use real life data and obtain a control chart equivalent to Shewhart $\bar{X}$ control chart, and yet make decisions within a reasonable short time acceptable to manufacturers? The first attempt to consider this question was made by Gumbel

and Littauer [5] who considered the first failures of product within small samples when product life is normally distributed. Control limits for sample minima (or sample maxima) were developed considering the assymmetric form of the extremal distribution. Factors for estimating the mean and the standard deviation of the parent distribution from observations of the minima (or maxima) were also given.

In an earlier work Howell [6] considered control charts which make use of sample minima and maxima together. This procedure however, did not provide any advantage in life testing over Shewhart's procedure since it required observation of both minima and maxima of a sample. In addition, Howell failed to recognize the assymmetric form of the extremes of a normal distribution* and used symmetric control limits for his control chart.

The exponential distribution has been widely used in describing equipment life in many studies on life testing and reliability [1]. The time saving feature of observing the rth failure in samples of size N, $(N > r)$, has been demonstrated by Epstein and Sobel [3] for the exponential distribution. Jacobson [7] has shown that a test procedure based on the average of lowest three out of five observations gives almost the same operating characteristic curve as in the test based on the average of five out of five and four out of four observations from a normal distribution.

In this paper we shall consider only exponentially distributed life data, and by using only the times to first failure in samples of size N, we shall establish four different control charts based on four different statistics of sample minima. These control charts will be compared with equivalent Shewhart $\bar{X}$ control charts in terms of time, cost, and ability to detect changes in the mean life of the manufactured product.

II. SOME PROPERTIES OF THE EXTREMES OF THE EXPONENTIAL DISTRIBUTION, AND THEIR USE IN LIFE TESTING

Let us consider a process where a product is being manufactured. Product life X is assumed to be exponentially distributed with a density function of the form

$$f(x;\theta) = \frac{1}{\theta} e^{-x/\theta}, \qquad x \geq 0, \qquad \theta > 0. \tag{1}$$

The parameter θ is the unknown mean life of this product. The purpose of the control chart procedure is first, to establish (under the assumption of statistical stability) a range of variation of the product life which will be due to chance causes; and second, by continuing observation of the product life to detect the outages which may be due to assignable causes requiring corrective action. Since θ is the only parameter in our parent distribution a "good" estimator of θ will suffice for construction of control limits.

Let $X_1, X_2, \ldots, X_N$ denote the failure times of N items of a product being tested simultaneously. It can be shown that

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2}$$

*See Gumbel [4], pp. 129–131.

is a sufficient, efficient, unbiased, and maximum likelihood estimator of θ [3]. The time required to obtain $\hat{\theta}$ is

$$Y = \max(X_1, X_2, \ldots, X_N); \tag{3}$$

and the time to observe the first failure is

$$Z = \min(X_1, X_2, \ldots, X_N). \tag{4}$$

The random variables Y and Z defined in (3) and (4), respectively, will have the following well known extremal distribution functions

$$F_Z(z;\theta,N) = 1 - e^{-zN/\theta}, \qquad z \geq 0, \qquad \theta > 0 \tag{5}$$

$$F_Y(y;\theta,N) = (1 - e^{-y/\theta})^N, \qquad y \geq 0, \qquad \theta > 0. \tag{6}$$

The expected time $E[Z]$ for the first failure and the expected time $E[Y]$ for the last failure will then be (See [4] for the development of (8))

$$E[Z] = \theta/N \tag{7}$$

$$E[Y] = \theta \sum_{\nu=1}^{N} 1/\nu. \tag{8}$$

Thus,

$$E[Y] - E[Z] = \theta \left[\sum_{\nu=1}^{N} 1/\nu - 1/N \right] = \theta \sum_{\nu=1}^{N-1} 1/\nu > 0. \tag{9}$$

Considering this time advantage one may propose to run M consecutive experiments on samples of size N and wait only for the first failures to occur instead of waiting for the last failure of one sample of size N. Thus, the new estimator of θ can be

$$\hat{\theta} = N/M \sum_{i=1}^{M} z_i, \tag{10}$$

where z_i is the first failure time in ith sample of size N as defined in (4). This estimator has the same statistical properties as the one given by (2).* Duration of an experiment to obtain this estimator will be

$$d_Z = \sum_{i=1}^{M} z_i, \tag{11}$$

*See appendix for the properties of this estimator.

and the expected duration $E[D_Z]$ will be

$$E[D_Z] = M\theta/N. \tag{12}$$

An alternative way of obtaining the average minimum in (10) could be to run M experiments on samples of size N simultaneously. In this case the time to obtain all of the M first failures will be:

$$D'_z = \text{Max}\ (z_1, z_2, \ldots, z_M). \tag{13}$$

This new random variable D'_Z, will have a distribution function of the form

$$F_{D'Z}(d'_Z; \theta, M, N) = [1 - e^{-d'_Z N/\theta}]^M, \qquad d'_Z \geq 0, \qquad \theta > 0 \tag{14}$$

and

$$E[D'_z] = \frac{\theta}{N} \sum_{\nu=1}^{M} 1/\nu. \tag{15}$$

It is clear from the above result that it is possible to estimate θ within a time much shorter than the time for observing the last failure of a sample of size N, by observing the M first failures of independent, identically distributed M samples.

We shall now develop four different statistics with first failures and construct specific control charts for each.

III. CONTROL CHARTS FOR MINIMUM LIFE DATA

Let us consider a continuing sampling procedure where a number of samples of size N are being tested. A test is terminated upon observation of the first failure. Thus, given θ, the mean product life, the duration of each test will depend only on N. After M samples are tested the results provide a sample of size M of the first failure times for interpretation. We shall consider the following four statistics for constructing control charts:

i) First failure time in samples of size N.
ii) Smallest of M first failure times in M samples of size N.
iii) Largest of M first failure times in M samples of size N.
iv) Average of M first failure times in M samples of size N.

1. Control Chart for the First Failure Times in Samples of Size N

We consider the random variable X, lifetime of a manufactured product with a density function of the form given in (1). Let $x_1, x_2, \ldots, x_N$ denote the life times of N specimens in a sample. We are interested in constructing control charts by using only the smallest of $x_1, x_2, \ldots, x_N$, i.e., the first failure time z which is defined in (4). Under the assumption of statistical stability the value of $\hat{\theta}$ can be estimated by observing M values of the random variable Z by using the estimator $\hat{\theta}$ given in (10).

If this process is under statistical control, future observations of Z will fall within two established

limits, namely upper control limit (UCL_Z) and lower control limit (LCL_Z) with a probability α. Thus

$$\int_{LCL_Z}^{UCL_Z} f(t)\,dt = \alpha. \tag{16}$$

We may express this relationship in the following form to facilitate a solution for UCL_Z and LCL_Z.

$$\int_0^{LCL_Z} f_Z(t)\,dt = \int_{UCL_Z}^{\infty} f_Z(t)\,dt = \frac{1-\alpha}{2}; \tag{17}$$

using (5) will lead to the solutions

$$UCL_Z = -\frac{\hat{\theta}}{N} \ln\left(\frac{1-\alpha}{2}\right) \tag{18}$$

$$LCL_Z = -\frac{\hat{\theta}}{N} \ln\left(\frac{1+\alpha}{2}\right). \tag{19}$$

Since $\hat{\theta}$ can be expressed as

$$\hat{\theta} = \frac{N}{M} \sum_{i=1}^{M} z_i = N\bar{z}, \tag{20}$$

we shall have

$$UCL_Z = -\bar{z} \ln\left(\frac{1-\alpha}{2}\right) \tag{21}$$

$$LCL_Z = -\bar{z} \ln\left(\frac{1+\alpha}{2}\right). \tag{22}$$

These limits can be considered as the upper and lower bounds of an acceptance region in testing the null hypothesis $\theta=\theta_1$ against the alternative $\theta=\theta_2$ with a type I error of $(1-\alpha)$. We note that the power of the test will be a function of θ_1/θ_2, where θ_1 is the mean life in the null hypothesis and θ_2 is the mean life in the alternate hypothesis. Operating characteristic curves for this test as functions of θ_1/θ_2 ratio can be obtained from

$$\text{Prob}\left\{-\frac{\theta_1}{N} \ln\left(\frac{1+\alpha}{2}\right) \leq Z \leq -\frac{\theta_1}{N} \ln\left(\frac{1-\alpha}{N}\right) \mid \theta = \theta_2\right\}, \tag{23}$$

which leads to the expression

$$\text{Prob}\left\{\text{Accept } H_0 \mid H_1\right\} = \left(\frac{1+\alpha}{2}\right)^{\theta_1/\theta_2} - \left(\frac{1-\alpha}{2}\right)^{\theta_1/\theta_2}. \tag{24}$$

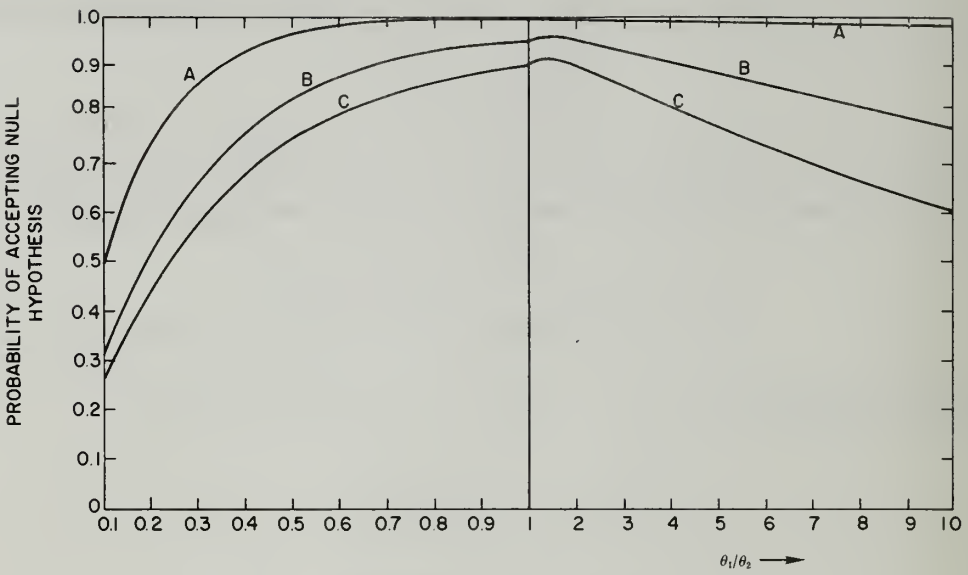

FIGURE 1. O. C. curves for the control chart of the minimums and the minimum of minima for $\alpha = 0.9973$ (A); $\alpha = 0.95$ (B); $\alpha = 0.90$ (C).

These curves are shown in Figure 1 for $\alpha = 0.9973$ which corresponds to the percentiles equivalent to 3σ limits of Shewhart $\bar{X}$ control chart.

2. Control Charts for the Smallest of the M First Failure Times in Samples of Size N

In this case we observe in M samples M first failures and consider the smallest failure time among M. Thus

$$\eta_{N,M} = \text{Min}\,(z_1, \ldots, z_M). \tag{25}$$

The random variable $\eta_{N,M}$ will have a density function of the form

$$f_\eta(t) = \frac{NM}{\theta} e^{-NMt/\theta} \qquad t \geq 0, \qquad \theta > 0, \tag{26}$$

which is the same as the density function of the random variable Z with, however, a different mean. Following the same argument as in the previous case, one can obtain the following control limits:

$$\text{UCL}_\eta = -\frac{\bar{z}}{M} \ln\left(\frac{1-\alpha}{2}\right) \tag{27}$$

and

$$LCL_\eta = -\frac{\bar{z}}{M} \ln\left(\frac{1+\alpha}{2}\right) \tag{28}$$

and similarly the expression for OC curves will be

$$\text{Prob}\left\{\text{Accept } H_0 \mid H_1\right\} = \left(\frac{1+\alpha}{2}\right)^{\theta_1/\theta_2} - \left(\frac{1-\alpha}{2}\right)^{\theta_1/\theta_2} \tag{29}$$

Note that even though we have different control limits the performance of this control chart is identical to that of the previous one.

3. Control Charts for the Largest of M First Failures in Samples of Size N

Let us consider the sample maxima of M minimums defined by

$$U_{N,M} = \text{Max } (Z_1, \ldots, Z_M). \tag{30}$$

The random variable $U_{N,M}$ will have a density function of the form

$$f_{U_{N,M}}(t) = \frac{MN}{\theta} e^{-Nt/\theta}(1-e^{-Nt/\theta})^{M-1} \quad t \geq 0,\ \theta > 0. \tag{31}$$

It is known that the mean can be expressed by

$$E(U_{NM}) = \frac{\theta}{N}\sum_{\nu=1}^{M} 1/\nu. \tag{32}$$

By evaluating the relationship given in (17) we shall obtain

$$LCL_U = -\frac{\hat{\theta}}{N}\ln\left[1-\left(\frac{1-\alpha}{2}\right)^{1/M}\right] \tag{33}$$

and

$$UCL_U = -\frac{\hat{\theta}}{N}\ln\left[1-\left(\frac{1+\alpha}{2}\right)^{1/M}\right] \tag{34}$$

or, by using $\hat{\theta}$ given in (20), we shall have

$$LCL_U = -\bar{z}\ln\left[1-\left(\frac{1-\alpha}{2}\right)^{1/M}\right] \tag{35}$$

and

$$UCL_U = -\bar{z}\ln\left[1-\left(\frac{1+\alpha}{2}\right)^{1/M}\right]. \tag{36}$$

The OC curves for this control chart can be obtained from

$$\text{Prob}\left\{-\frac{\theta_1}{N}\ln\left[1-\left(\frac{1-\alpha}{2}\right)^{1/M}\right] \leq U_{N,M} \leq \frac{-\theta_1}{N}\ln\left[1-\left(\frac{1+\alpha}{2}\right)^{1/M}\right] \mid \theta = \theta_2\right\}, \tag{37}$$

which will lead to the expression

$$\text{Prob}\{\text{Accept } H_0|H_1\}=\left\{1-\left[1-\left(\frac{1+\alpha}{2}\right)^{1/M}\right]^{\theta_1/\theta_2}\right\}^M-\left\{1-\left[1-\left(\frac{1-\alpha}{2}\right)^{1/M}\right]^{\theta_1/\theta_2}\right\}^M \tag{38}$$

These curves for several values of M and for $\alpha=0.9973$ are shown in Figure 2.

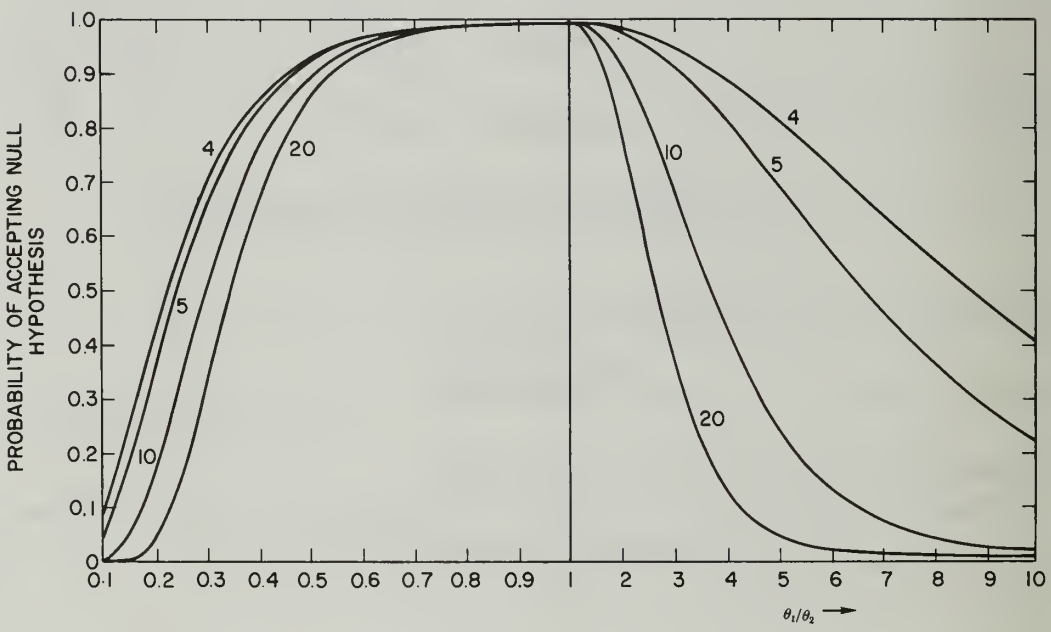

FIGURE 2. O. C. curves for the control chart of the maximum of minima for $\alpha=0.9973$ and $M=4, 5, 10$, and 20.

4. Control Charts for the Average of M First Failure Times in Samples of Size N

Consider the statistic

$$V=\frac{1}{M}\sum_{i=1}^{M} z_i=\bar{z}, \tag{39}$$

where z_i are the first failure times in samples of size N. The random variable V has the same distribution as $\hat{\theta}/N$. Since

$$f_{\hat{\theta}}(t)=\frac{M/\theta(Mt/\theta)^{M-1}e^{-Mt/\theta}}{(M-1)!}, \qquad t\geq 0 \qquad \theta>0 \tag{40}$$

the density function of V will be

$$f_V(t)=\frac{MN/\theta(MNt/\theta)^{M-1}e^{-MNt/\theta}}{(M-1)!}, \qquad t\geq 0, \qquad \theta>0 \tag{41}$$

and

$$E[V]=\theta/N, \qquad \text{Var}\,[V]=\theta^2/MN^2. \tag{42}$$

Upper and lower control limits can be obtained by using (41) in (17). Since $2M\hat{\theta}/\theta$ has a Chi-Square distribution with $2M$ degrees of freedom and $V=\hat{\theta}/N$, control limits will be

$$UCL_V=\frac{\hat{\theta}}{2MN}\chi^2_{2M,\frac{1+\alpha}{2}} \tag{43}$$

$$LCL_V=\frac{\hat{\theta}}{2MN}\chi^2_{2M,\frac{1-\alpha}{2}} \tag{44}$$

or by using (10) we shall obtain

$$UCL_V=\frac{\bar{z}}{2M}\chi^2_{2M,\frac{1+\alpha}{2}} \tag{45}$$

$$LCL_V=\frac{\bar{z}}{2M}\chi^2_{2M,\frac{1-\alpha}{2}} \tag{46}$$

where $\chi^2_{2M,\gamma}$ denotes the γ percentile of a Chi-Square distribution with $2M$ degrees of freedom.

Operating characteristic curves for this control chart can be obtained from

$$\text{Prob}\left\{\frac{\theta_1}{2MN}\chi^2_{2M,\frac{1-\alpha}{2}}\leq V\leq\frac{\theta_1}{2MN}\chi^2_{2M,\frac{1+\alpha}{2}}\;\middle|\;\theta=\theta_2\right\} \tag{47}$$

or

$$\text{Prob}\left\{\frac{\theta_1}{\theta_2}\chi^2_{2M,\frac{1-\alpha}{2}}\leq\frac{2NM}{\theta_2}V\leq\frac{\theta_1}{\theta_2}\chi^2_{2M,\frac{1+\alpha}{2}}\right\}, \tag{48}$$

where $\frac{2NM}{\theta_2}V$ has a Chi-Square distribution with $2M$ degrees of freedom. Operating characteristic curves for $\alpha=0.9973$ and for several values of M are shown in Figure 3.

IV. COMPARISON WITH SHEWHART $\bar{X}$ CONTROL CHARTS

In order to compare these four new control charts with the Shewhart $\bar{X}$ control chart, we shall first obtain OC curves for $\bar{X}$ control chart for the case when product life X has an exponential distribution. According to Shewhart $\bar{X}$ control chart procedure UCL and LCL are given by

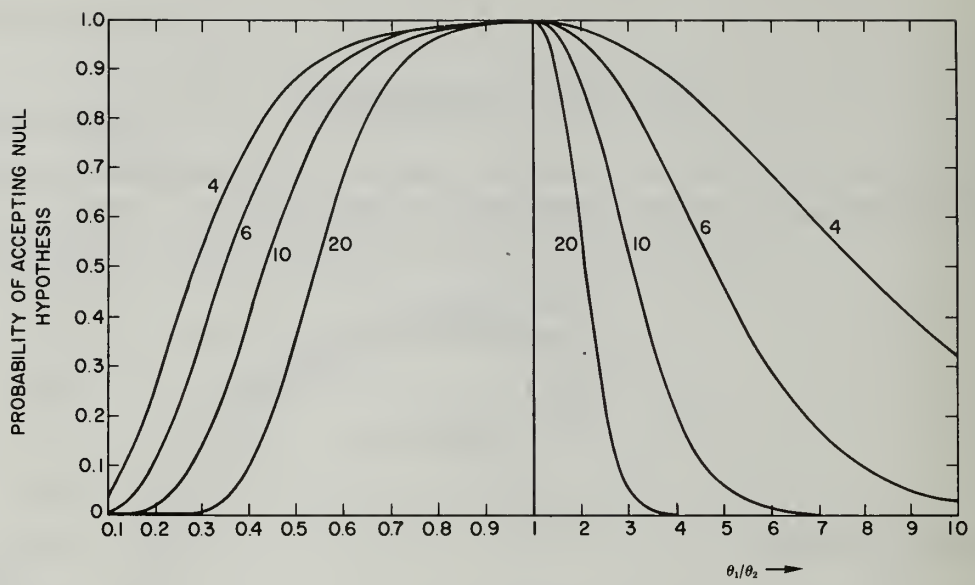

FIGURE 3. O. C. curves for the control chart of the average of minima for $\alpha = 0.9973$ and $M = 4, 6, 10$, and 20.

$$UCL = \bar{\bar{X}} + 3\frac{S}{\sqrt{N}} \tag{49}$$

$$LCL = \bar{\bar{X}} - 3\frac{S}{\sqrt{N}} \tag{50}$$

where $\bar{\bar{X}}$ and S are the estimators of the mean and the standard deviation of the parent distribution. Since X is exponentially distributed, the mean and standard deviation are equal to θ. Then the relationships given in (49) and (50) will become

$$UCL = \theta\left(1 + \frac{3}{\sqrt{N}}\right) \tag{51}$$

$$LCL = \theta\left(1 - \frac{3}{\sqrt{N}}\right). \tag{52}$$

This leads to the expression

$$\text{Prob}\left\{\theta\left(1 - \frac{3}{\sqrt{N}}\right) \leq \bar{X} \leq \theta\left(1 + \frac{3}{\sqrt{N}}\right)\right\} = \alpha. \tag{53}$$

Operating characteristic curves can now be constructed from

$$\text{Prob}\left\{\theta_1\left(1 - \frac{3}{\sqrt{N}}\right) \leq \bar{X} \leq \theta_1\left(1 + \frac{3}{\sqrt{N}}\right) \,\middle|\, \theta = \theta_2\right\} \tag{54}$$

This is equivalent to

$$\text{Prob}\left\{\frac{\theta_1}{\theta_2}\left(1-\frac{3}{\sqrt{N}}\right)2N \leq \bar{X}\frac{2N}{\theta_2} \leq \frac{\theta_1}{\theta_2}\left(1+\frac{3}{\sqrt{N}}\right)2N\right\}. \tag{55}$$

Noting that $\bar{X}2N/\theta_2$ has a Chi-Square distribution with $2N$ degrees of freedom we obtain

$$\text{Prob}\left\{2\frac{\theta_1}{\theta_2}(N-3\sqrt{N}) \leq \chi^2_{2N} \leq 2\frac{\theta_1}{\theta_2}(N+3\sqrt{N})\right\}. \tag{56}$$

These curves are shown for different values of N in Figure 4.

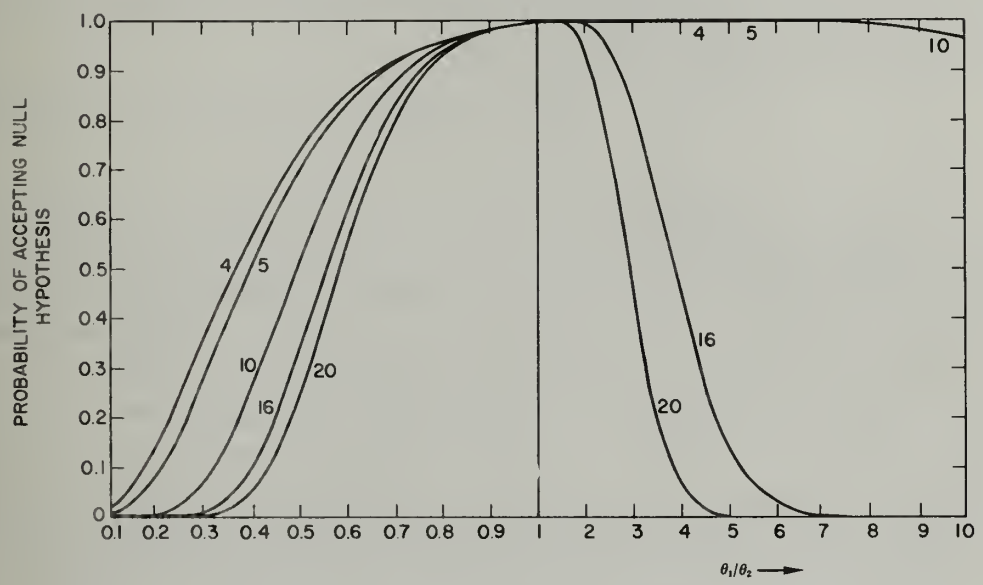

FIGURE 4. O. C. curves for the Shewhart $\bar{x}$ control chart (when the random variable X is exponentially distributed) for $\alpha = 0.9973$ and $M = 4, 5, 10, 16$, and 20.

1. Comparison Based on Equal Testing Duration

In Section II we have shown that the expected duration of a life test in a sample of size N is

$$E[Y] = \theta \sum_{\nu=1}^{N} 1/\nu, \tag{8}$$

if all N values of failure times are observed. This represents the expected time required to obtain one point on Shewhart $\bar{X}$ chart. On the other hand the expected time to obtain one point of an *extremal* control chart which uses only the first failure data is

$$E[Z] = \theta/N \tag{7}$$

for the minima chart; and

$$E[D_Z]=\frac{M\theta}{N} \tag{12}$$

for the control charts on minima of the minima, maxima of the minima and the average of minima. Where M represents the number of samples of size N used to obtain the statistics of minima. We have also expressed the expected duration of such an experiment if M samples were tested simultaneously by

$$E[D_Z]=\frac{\theta}{N}\sum_{\nu=1}^{M} 1/\nu. \tag{15}$$

Let us now consider a Shewhart type $\bar{X}$ control chart versus a control chart based on the minimums of samples of size N. By equating (7) and (8) we obtain

$$\sum_{\nu=1}^{N_1} 1/\nu = 1/N_2, \tag{57}$$

where N_1 is the sample size in Shewhart $\bar{X}$ control chart and N_2 is the sample size in the minima chart. Since the only solution to this above equation is $N_1=N_2=1$, we can conclude that by increasing N_2 one can decrease the duration of the minima experiment without affecting the OC curve (see Fig. 1), for the OC curve for this control chart is independent of N_2. In general, the performance of minima control chart is poorer than the Shewhart $\bar{X}$ control chart, expecially for $\theta_1/\theta_2 < 1$; although for $N_1 < 10$ the Shewhart $\bar{X}$ control chart's performance is equally as poor as the minima control chart for $\theta_1/\theta_2 > 1$.

In the case of a control chart on the minimum of the minima, one can use a similar argument. Considering the facts that the distribution of the minimum of minima has the same form as the distribution of the minima, except for a smaller mean, and that their OC curves are identical, a control chart on the minimum of minima will not offer any advantage over the control chart of the minima.

The control chart on the maximums of minima and the Shewhart $\bar{X}$ chart can be compared on the basis of equal duration. Equating the relationships given in (8) and (12) we obtain

$$\sum_{\nu=1}^{N_1} 1/\nu = \frac{M_2}{N_2}, \tag{58}$$

where N_1 is the sample size used in the Shewhart $\bar{X}$ control chart, N_2 is the sample size in the maximum of the minima control chart for observing a minimum, and M_2 is the sample size to observe a maximum among M_2 minima. Several feasible sets of N_1, N_2, M_2 can be obtained from the above relationship. Since the OC curves (see Figure 2) are independent of N_2, one can always select a feasible (or near feasible) pair (M_2, N_2) so that the probability of accepting the null hypothesis $(\theta=\theta_1)$ is smaller than in the case of Shewhart $\bar{X}$ charts. Furthermore, since durations of the tests are the same for the average of minima chart and the maximum of minima chart, the latter chart has the same advantage over the Shewhart $\bar{X}$ control chart. Furthermore, as one can see immediately from Figures 2 and 3 that for the same pairs of N and M OC curves of the average of minima control chart indicate clearly the advantage of using the average of minima control chart instead of the maxima of the minima chart.

2. Comparison Based on Costs

In order to compare these control charts on the basis of cost, we shall first formulate the cost of obtaining one point on the control chart for each case.

Let

C_d = The cost of one specimen if it is destroyed during the life test.

V_s = The reduction in the value of one specimen if it is subjected to the life test, but survived the test.

$C_t(N)$ = Total cost per unit time for running a life test with N specimens.

$C_0(N)$ = Set-up cost for a life test on a sample of size N.

The expected cost of obtaining one point on the Shewhart $\bar{X}$ chart is

$$C_1 = C_0(N_1) + N_1 C_d + C_t(N_1)\theta \sum_{\nu=1}^{N_1} 1/\nu. \tag{59}$$

The expected cost of obtaining one point on the minima control chart is

$$C_2 = C_0(N_2) + C_d + (N_2 - 1)V_s + C_t(N_2)\frac{\theta}{N_2} \tag{60}$$

and the expected cost of obtaining one point on the minima of minima, maximum of the minima and the average of minima chart is

$$C_3 = M_2 C_0(N_2) + M_2 C_d + M_2(N_2 - 1)V_s + C_t(N_2)\frac{M_2\theta}{N_2}. \tag{61}$$

One can immediately conclude from OC curves that for any choice of M_2 and N_2 the average of minima control chart will always give a smaller probability of acceptance than the minimum of the minima and maximum of the minima control charts. As for the minima control chart, since probability of acceptance is not a function of M_2 and N_2 one can always find a pair M_2^*, N_2^* for which the average of minima will yield a smaller probability of acceptance than the minima chart for the same cost.

We shall compare the Shewhart $\bar{X}$ control chart with the average of minimum control chart for the same probability of acceptance and the same θ_1/θ_2 ratio. By selecting such values on their respective OC curves, we can compute related costs. For example for $\theta_1/\theta_2 = 3$ and $N_1 = 16$ the probability of accepting the null hypothesis $(\theta = \theta_1)$ when actually $\theta = \theta_2$, will be approximately 0.8444, shown on the Shewhart $\bar{X}$ chart OC curves. For the same probability of acceptance and θ_1/θ_2 ratio the OC curves for the average of minima chart will give $M_2 \simeq 6$. Thus the related costs will be, respectively,

$$C_1 = C_0(16) + 16C_d + C_t(16)3.381\theta$$

and

$$C_3 = 6C_0(N_2) + 6C_d + 6(N_2 - 1)V_s + C_t(N_2)\frac{6\theta}{N_2},$$

whose difference is

$$C_1 - C_3 = C_0(16) - 6C_0(N_2) + 10C_d - 6(N_2 - 1)V_s + \theta[3.381C_t(16) - \frac{6}{N_2} C_t(N_2)].$$

Given the specific forms of $C_t(N)$ and $C_0(N)$, one can find the value of N_2 which maximizes this cost difference. For illustration purposes let us assume that

$$C_0(N_2) = A_0 N_2 \qquad A_0 > 0,\ N_2 \geq 1$$

and

$$C_t(N_2) = A_t N_2 \qquad A_t > 0,\ N_2 \geq 1.$$

Thus, the difference will be

$$C_1 - C_2 = -6(A_0 + V_s)N_2 + 16A_0 + 10C_d + 6V_s + 48.09\theta A_t.$$

This difference decreases linearly with N_2. Consequently, the maximum cost difference will be obtained by setting $N_2 = 1$, which will give

$$C_1 - C_2 = 10(A_0 + C_d) + 48.09A_t\theta.$$

It is possible to increase this difference further by using (15) instead of (12) in expressing the expected duration in C_3. Then the expected cost difference will become

$$C_1 - C_3 = 10(A_0 + C_d) + 51.65A_t\theta.$$

Note that this second difference reflects the cost advantage related to the use of the initial distribution (exponential) instead of a normal approximation for the average of a sample from exponential distribution.

APPENDIX

In the following we shall prove that the extimator $\hat{\theta}$ given in (10) is an unbiased, maximum likelihood, sufficient, efficient and minimum variance estimator of the mean of exponential distribution.

The random variable Z defined by (4) has an exponential distribution given in (5). The moment generating function of Z can be expressed by

$$MGF_z(t) = \int_0^\infty \frac{N}{\theta} e^{tz} e^{-zN/\theta}\, dz = \frac{N}{N - \theta t}. \tag{62}$$

Thus, the random variable $D_Z = \sum_{1=1}^{M} Z_i$ will have the following moment generating function

$$MGF_{D_Z}(t) = \left(\frac{N/\theta}{N/\theta - t}\right)^M. \tag{63}$$

Since $\hat{\theta}=N/M\,D_Z$, the moment generating function of $\hat{\theta}$ will be

$$MGF_{\hat{\theta}}(t)=\left(\frac{M/\theta}{M/\theta-t}\right)^M. \tag{64}$$

We note that this is the moment generating function of Erlang distribution of the form

$$G_{\hat{\theta}}(t;m,\theta)=\frac{M/\theta(M/(\theta t))^{M-1}e^{-Mt/\theta}}{(M-1)!},\qquad \frac{M}{\theta}>0,\qquad t\geq 0. \tag{65}$$

It follows from (65) that

$$E[\hat{\theta}]=\theta \qquad \text{and} \qquad \text{Var}\,[\hat{\theta}]=\frac{\theta^2}{M}. \tag{66}$$

Thus $\hat{\theta}$ is an unbiased estimator of θ.

It also follows from (65) that $2M/\theta\,\hat{\theta}$ has a Chi-Square distribution with $2M$ degrees of freedom, or the distribution of $\hat{\theta}$ is $\theta/2M\,\chi^2_{2M}$.

Let us consider the joint density

$$f(z_1,z_2,\ldots,z_M;\theta)=\left(\frac{N}{\theta}\right)^M e^{-\frac{N}{\theta}\sum_{i=1}^{M} z_i}. \tag{67}$$

The maximum likelihood estimator of θ can be found by differentiating (67) with respect to θ and equating it to zero.

$$\frac{\partial f(z_1,z_2,\ldots,z_M,\theta)}{\partial\theta}=\left[\frac{N}{\theta}\sum_{i=1}^{M} z_i-M\right]\frac{N^M}{\theta^{M+1}}e^{-\frac{N}{\theta}\sum_{i=1}^{M} z_i}=0 \tag{68}$$

The value of θ which satisfies (68) will be

$$\hat{\theta}=\frac{N}{M}\sum_{i=1}^{M} z_i.$$

Thus, $\hat{\theta}$ is also a maximum likelihood estimator.

The joint density function given in (67) can be expressed as

$$f(z_1,z_2,\ldots,z_M;\theta)=\left(\frac{N}{\theta}\right)^M e^{-\frac{M}{\theta}\hat{\theta}}=1\cdot h(\hat{\theta},\theta), \tag{69}$$

which indicates that $\hat{\theta}$ is also a sufficient estimator of θ.

Following Epstein and Sobel [3], we can show that $\hat{\theta}$ is also an efficient and minimum variance

estimator. The Cramer-Rao lower bound $1/E\ (\partial\ \ln f/\partial\theta)^2$, where f is given by (67) will be compared as follows:

(70) $$\frac{\partial\ \ln\ f(z_1,\ .\ .\ .,\ z_M;\ \theta)}{\partial\theta} = \frac{M}{\theta}\left[\frac{\hat{\theta}}{\theta} - 1\right].$$

Thus

(71) $$E\left[\frac{\partial\ \ln f}{\partial\theta}\right]^2 = E\left[\frac{M}{\theta}\left(\frac{\hat{\theta}}{\theta} - 1\right)\right]^2 = \frac{M}{\theta^2},$$

where

$$E(\hat{\theta}) = \theta,\ E(\hat{\theta}^2) = \text{Var}\ (\hat{\theta}) + \theta^2\ = \frac{\theta^2}{M} + \theta^2.$$

Consequently the lower bound will be

(72) $$1/E(\partial\ \ln\ f/\partial\theta)^2 = \theta^2/M.$$

Note that the variance of $\hat{\theta}$, given by (66) is equal to this lower bound. This means that the variance of any other estimate will be equal or greater than this bound. Thus $\hat{\theta}$ is a minimum variance estimator. The efficiency of an estimator is defined as the ratio between the lower bound of the variance and the variance of the estimator.* In this particular case the efficiency is equal to one i.e., $\hat{\theta}$ is also an efficient estimator.

REFERENCES

[1] Buckland, W. R., *Statistical Assessment of the Life Characteristic* (Hafner Publishing Company, 1964).

[2] Cramer, H., *Mathematical Methods of Statistics* (Princeton University Press, 1946).

[3] Epstein, B. and M. Sobel, "Life Testing," J. Am. Stat. Assn. *48*, 486–501 (1953).

[4] Gumbel, E. J., *Statistics of Extremes* (Columbia University Press, 1958), p. 115.

[5] Gumbel, E. J. and S. B. Littauer, "Statistical Control of Sample Extremes," Technical report, Operations research group Columbia University School of Engineering and Applied Science (Feb. 1967).

[6] Howell, J. M., "Control Chart for Largest and Smallest Values," Annals of Math. Stat. *20*, 305–309 (1949).

[7] Jacobson, P.H., "The relative power of three statistics for small sample destructive tests," J. Am. Stat. Assn. *42*, 575–584 (1947).

*See Cramer [2, p. 481] for this definition.

APPROXIMATE PROBABILITY DISTRIBUTIONS FOR THE EXTREME SPREAD

Malcolm S. Taylor and Frank E. Grubbs

U.S. Army Ballistic Research Laboratories
Aberdeen Proving Ground, Maryland

ABSTRACT

The extreme spread, or greatest distance between all pairs of impact points on a target, is often used as a rapid measure of dispersion or precision of shot groups on a target. It is therefore desirable to know its statistical properties. Since the exact theoretical distribution has not yet been worked out, this paper examines the accuracy of several approximations which are checked against large sample monte carlo values. We find in particular that for the sample sizes considered the extreme spread can be approximated well by a Chi variate.

. INTRODUCTION AND SUMMARY

In ballistics work, and especially in small arms firings at targets to determine accuracy, it is quite atural to measure the closeness of a group of shots by the "extreme spread," or the greatest distance etween any two shots of a group. Moreover, the extreme spread may be determined quickly with a uler and does not require any detailed or involved computation, as does the round-to-round standard eviation in each direction, or the mean radius of the shots of a group, for example. It is for these reasons iat ballisticians, riflemen, and others have long had a great interest in the extreme spread, for it is uly the most rapid measure of dispersion of shots on a target. Sometimes the extreme spread, or the iaximum distance between pairs of points on the target, is called the "group diameter," but there is very subtle difference between the two when one delves into the general problem on a statistical asis. We do not intend to cover all the pertinent details relating to the statistical analysis of patterns f shots on a target here, but interested readers might well study the booklet of Grubbs [4]. Rather, e intend to develop in this paper the properties of the extreme spread more extensively than has been one in the past, and thereby contribute to an improved understanding of the statistical characteristics f the probability distribution of the extreme spread, which is required in any first-class or overall nalysis of target accuracy studies.

In our introduction of the subject, we point out that the extreme spread is a random variable hich follows some kind of statistical or probability distribution. Indeed, the amount of random variation rom one group of shots to another depends markedly on the sample size, or the number of shots in a roup, and the underlying unknown, population round-to-round standard deviation, which we will call . The population standard deviation, σ, is a one-directional or "linear" quantity, say, for the x or orizontal direction (as well as the y or vertical direction), and for a very large number of shots it may e found as the square root of the sum of squares of deviations in the x-direction from the mean divided y the number of rounds. In rifle firing, and in many other types of weapon studies, the population

standard deviations in the two directions are equal or very nearly so. Hence, it may be assumed in our following analysis that $\sigma_x = \sigma_y = \sigma$. The exact theoretical probability distribution of the extreme spread, or bivariate range, as it is often called, has not been determined as of this date, although many of the key properties of the distribution are fairly accurately known from previous studies. In the following, we report on the results of a Monte Carlo type of computer simulation, along with the necessary statistical analyses, to find approximate statistical distributions which for all the practical purposes result in the degree of accuracy needed to round out sufficiently our understanding of properties of the distribution of the extreme spread, at least for the very important practical cases involving small sample sizes of predominant interest.

Our acknowledgements must go to Mr. Philip G. Rust, retired industrialist of the Winnstead Plantation, Thomasville, Georgia, for his great interest in critical analyses of accuracy firings of rifles, which provided much of the motivation for this investigation, as well as for the booklet by Grubbs [4], which are of importance to ballistic analyses generally.

II. SOME ANALYTICAL PRELIMINARIES

Consider a random sample $X_1, \ldots, X_n$ $(X_i = (x_i, y_i))$ from a bivariate normal distribution with probability density function (p.d.f.) given by

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma}; \qquad (\sigma_x = \sigma_y = \sigma). \tag{1}$$

The extreme spread (ES), or bivariate range, is defined as $\mathrm{ES} = \max_{i,j} |X_i - X_j|$. This ES is of course a random variable, as previously pointed out, and we seek its probability distribution, realizing that it will be dependent upon the sample size n. For the case $n = 2$, for example, we have $ES = |X_1 - X_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \simeq 2\chi\sigma$, where the random variate chi has two degrees of freedom. In this case, the mean value of the extreme spread, $E(ES) = 1.77245\sigma$, and the variance $V(ES) = (0.9265\sigma)^2$. This result does not have a direct extension to higher sample sizes, $(n > 2)$ however, and the distribution of the extreme spread ES has not been determined analytically.

Some earlier work of Wilks and Grubbs in Reference [4] have led to Monte Carlo estimates of the first four moments of the three-dimensional or trivariate range, the trivariate midrange, the extreme spread or bivariate range, and the bivariate midrange. Cacoullos and DeCicco [2] have investigated two approximations of the ES distribution based on the Monte Carlo moment data of Wilks and Grubbs. In this paper we present improved Monte Carlo moment estimates (by virtue of greatly increased sample size), as well as quantile estimates, or percentage points, which were not available previously and propose some approximate distributions, which will suffice for most analytical studies in practice. The large-sample moments tabulated below include the mean μ, standard deviation σ, the skewness measure α_3, and "kurtosis" or peakedness measure α_4. (See any standard textbook on statistics for further definitions and formulas.)

III. MONTE CARLO RESULTS

The moments (Table 1) and various quantiles of interest (Table 2) are based upon 10^4 Monte Carlo samples of the random variable ES for each value of sample size, n. The value of n specifies the number

TABLE 1. *Moment Constants of the Extreme Spread*

n	μ_{ES}	σ_{ES}	α_3	α_4	n	μ_{ES}	σ_{ES}	α_3	α_4
2	1.766	0.932	0.632	3.294	10	3.813	0.745	0.388	3.288
3	2.406	0.887	0.451	3.143	15	4.190	0.694	0.395	3.255
4	2.787	0.856	0.393	3.163	20	4.452	0.668	0.400	3.240
5	3.066	0.828	0.390	3.171	25	4.639	0.650	0.439	3.307
6	3.277	0.806	0.374	3.194	28	4.734	0.642	0.426	3.357
7	3.443	0.783	0.373	3.177	30	4.788	0.635	0.463	3.441
8	3.582	0.771	0.392	3.231	31	4.822	0.631	0.434	3.321
9	3.710	0.754	0.382	3.215	34	4.891	0.623	0.422	3.318

NOTE: The numbers in the second column are $E(ES)/\sigma$, and those of the third column are $SD(ES)/\sigma$.

TABLE 2. *Percentage Points of the Extreme Spread*

n	$P_{0.0005}$	$P_{0.0010}$	$P_{0.0050}$	$P_{0.0100}$	$P_{0.0250}$	$P_{0.0500}$	$P_{0.1000}$	$P_{0.9000}$	$P_{0.9500}$	$P_{0.9750}$	$P_{0.9900}$	$P_{0.9950}$	$P_{0.9990}$	$P_{0.9995}$
3	0.339	0.383	0.578	0.687	0.882	1.066	1.313	3.588	3.984	4.318	4.746	5.002	5.595	5.834
4	0.653	0.710	0.946	1.076	1.283	1.477	1.725	3.916	4.285	4.602	5.010	5.290	5.938	6.190
5	0.885	0.983	1.260	1.400	1.611	1.801	2.046	4.156	4.519	4.832	5.207	5.461	6.057	6.288
6	1.137	1.227	1.491	1.636	1.853	2.043	2.278	4.336	4.670	4.973	5.361	5.655	6.221	6.431
7	1.348	1.452	1.710	1.842	2.043	2.243	2.477	4.480	4.805	5.110	5.471	5.728	6.245	6.427
8	1.525	1.608	1.863	1.998	2.208	2.403	2.636	4.595	4.937	5.227	5.582	5.848	6.379	6.621
9	1.607	1.709	2.030	2.167	2.373	2.563	2.786	4.702	5.029	5.308	5.672	5.930	6.398	6.658
10	1.798	1.884	2.140	2.277	2.482	2.669	2.896	4.786	5.118	5.409	5.750	6.004	6.552	6.742
15	2.295	2.372	2.656	2.772	2.963	3.129	3.340	5.101	5.396	5.668	6.000	6.235	6.727	6.897
20	2.630	2.721	2.972	3.095	3.276	3.438	3.626	5.336	5.630	5.880	6.205	6.436	6.890	6.998
25	2.894	2.965	3.220	3.329	3.504	3.652	3.845	5.494	5.790	6.049	6.364	6.578	7.012	7.198
28	2.952	3.044	3.312	3.424	3.605	3.759	3.953	5.575	5.860	6.113	6.453	6.664	7.138	7.323
30	3.084	3.170	3.402	3.511	3.678	3.834	4.017	5.619	5.898	6.170	6.476	6.711	7.205	7.386
31	3.149	3.216	3.429	3.541	3.712	3.868	4.055	5.651	5.927	6.180	6.503	6.719	7.146	7.317
34	3.216	3.297	3.517	3.630	3.797	3.946	4.127	5.706	5.979	6.224	6.523	6.731	7.218	7.389

of points sampled from the circular normal distribution to determine a single value of ES.

The quantiles of Table 2 are, of course, subject to standard definitions and interpretations. For example, for a sample size of $n=7$ and known population standard deviation σ, the lower 1-percent point is 1.842σ and therefore in random sampling from a bivariate normal population with standard deviation σ we would expect that only 1-percent of the extreme spreads for a sample of size seven would fall below 1.842σ. Similarly, for the 99-percent point, $P_{0.9900}$, (or upper 1-percent significance level) we would expect only 1-percent of the extreme spreads for sample size $n=10$ to exceed 5.75σ. The mean value of the extreme spread for a sample of size 10 from Table 1 is 3.813σ.

IV. APPROXIMATE PROBABILITY DENSITIES

Candidates for approximating the probability distribution of the extreme spread were chosen from three well known families of distributions, specifically, the chi distribution, the lognormal distribution, and the two-parameter Weibull distribution. Some particulars of our findings are detailed in the following paragraphs.

Chi Distribution Approximation

In considering the chi distribution, we made use of the fact that $2mu_n^2/v$ is approximately distributed as χ^2 with $\nu = 2m^2/v$ degrees of freedom (see, for example, Grubbs et al. [5]), where w_n is the sample range for random samples of size n from a *univariate* normal population. In our notation, $m = E(w_n^2)$ and $v = V(w_n^2)$. We used this same type of approximation for the extreme spread, ES, and interpolated linearly to evaluate chi-square for fractional degrees of freedom. Using the sample moments of ES to estimate m and v, this family provided a rather good fit to the sample quantiles over the entire range of n considered, although the dimension of the sample space has increased from one to two. The results are summarized in Table 3 where the italicized value is the fitted value juxtaposed to the Monte Carlo quantile estimate.

TABLE 3. *Chi Approximation*

	$P_{0.0050}$	$P_{0.0100}$	$P_{0.0250}$	$P_{0.0500}$	$P_{0.1000}$	$P_{0.9000}$	$P_{0.9500}$	$P_{0.9750}$	$P_{0.9900}$	$P_{0.9950}$	$P_{0.9990}$
3	0.578	0.687	0.882	1.066	1.313	3.588	3.984	4.318	4.746	5.002	5.595
	0.562	*0.676*	*0.868*	*1.057*	*1.297*	*3.593*	*3.975*	*4.306*	*4.717*	*4.997*	*5.578*
4	0.946	1.076	1.283	1.477	1.725	3.916	4.285	4.602	5.010	5.290	5.938
	0.905	*1.038*	*1.254*	*1.458*	*1.707*	*3.923*	*4.289*	*4.597*	*4.981*	*5.233*	*5.786*
5	1.260	1.400	1.611	1.801	2.046	4.156	4.519	4.832	5.207	5.461	6.057
	1.182	*1.325*	*1.550*	*1.758*	*2.011*	*4.164*	*4.516*	*4.813*	*5.177*	*5.424*	*5.939*
6	1.491	1.636	1.853	2.043	2.278	4.336	4.670	4.973	5.361	5.655	6.221
	1.410	*1.560*	*1.787*	*1.995*	*2.249*	*4.352*	*4.677*	*4.966*	*5.318*	*5.559*	*6.052*
7	1.710	1.842	2.043	2.243	2.477	4.480	4.805	5.110	5.471	5.728	6.245
	1.608	*1.756*	*1.984*	*2.191*	*2.442*	*4.483*	*4.799*	*5.083*	*5.417*	*5.647*	*6.127*
8	1.863	1.998	2.208	2.403	2.636	4.595	4.937	5.227	5.582	5.848	6.379
	1.755	*1.904*	*2.136*	*2.341*	*2.592*	*4.608*	*4.921*	*5.196*	*5.521*	*5.752*	*6.222*
9	2.030	2.167	2.373	2.563	2.786	4.702	5.029	5.308	5.672	5.930	6.398
	1.910	*2.060*	*2.288*	*2.495*	*2.740*	*4.704*	*5.012*	*5.280*	*5.599*	*5.820*	*6.277*
10	2.140	2.277	2.482	2.669	2.896	4.786	5.118	5.409	5.750	6.004	6.552
	2.027	*2.176*	*2.404*	*2.609*	*2.853*	*4.799*	*5.105*	*5.363*	*5.680*	*5.893*	*6.343*
15	2.656	2.772	2.963	3.129	3.340	5.101	5.396	5.668	6.000	6.235	6.727
	2.496	*2.642*	*2.864*	*3.061*	*3.299*	*5.109*	*5.387*	*5.627*	*5.915*	*6.115*	*6.524*
20	2.972	3.095	3.276	3.438	3.626	5.336	5.630	5.880	6.205	6.436	6.890
	2.808	*2.953*	*3.173*	*3.359*	*3.584*	*5.333*	*5.598*	*5.833*	*6.105*	*6.293*	*6.688*
25	3.220	3.329	3.504	3.652	3.845	5.494	5.790	6.049	6.364	6.578	7.012
	3.027	*3.171*	*3.383*	*3.571*	*3.800*	*5.499*	*5.758*	*5.980*	*6.246*	*6.429*	*6.809*
28	3.312	3.424	3.605	3.759	3.953	5.575	5.860	6.113	6.453	6.664	7.138
	3.144	*3.284*	*3.493*	*3.686*	*3.903*	*5.582*	*5.835*	*6.056*	*6.317*	*6.497*	*6.875*
30	3.402	3.511	3.678	3.834	4.017	5.619	5.898	6.170	6.476	6.711	7.205
	3.209	*3.351*	*3.560*	*3.746*	*3.961*	*5.626*	*5.877*	*6.100*	*6.359*	*6.533*	*6.906*
31	3.429	3.541	3.712	3.868	4.055	5.651	5.927	6.180	6.503	6.719	7.146
	3.256	*3.398*	*3.601*	*3.785*	*4.003*	*5.654*	*5.902*	*6.122*	*6.378*	*6.553*	*6.920*
34	3.517	3.630	3.797	3.946	4.127	5.706	5.979	6.224	6.523	6.731	7.218
	3.342	*3.485*	*3.688*	*3.871*	*4.085*	*5.715*	*5.959*	*6.173*	*6.424*	*6.598*	*6.958*

Lognormal Distribution Approximation

The lognormal distribution provides an excellent fit for large values of n. Following recommendations of Aitchison and Brown [1], we used the method of quantiles (specifically, the 10th and 90th percentiles) for purposes of estimating μ and σ of the associated normal distribution. Possibly, a

TABLE 4. *Log Normal Approximation*

n	$P_{0.0005}$	$P_{0.0010}$	$P_{0.0050}$	$P_{0.0100}$	$P_{0.0250}$	$P_{0.0500}$	$P_{0.1000}$	$P_{0.9000}$	$P_{0.9500}$	$P_{0.9750}$	$P_{0.9900}$	$P_{0.9950}$	$P_{0.9990}$	$P_{0.9995}$
15	2.295	2.372	2.656	2.772	2.963	3.129	3.340	5.101	5.396	5.668	6.000	6.235	6.727	6.897
	2.397	*2.478*	*2.697*	*2.811*	*2.986*	*3.146*	*3.340*	*5.101*	*5.416*	*5.706*	*6.062*	*6.317*	*6.877*	*7.108*
20	2.630	2.721	2.972	3.095	3.276	3.438	3.626	5.336	5.630	5.880	6.205	6.436	6.890	6.998
	2.679	*2.761*	*2.984*	*3.098*	*3.274*	*3.433*	*3.626*	*5.336*	*5.636*	*5.910*	*6.246*	*6.485*	*7.008*	*7.222*
25	2.894	2.965	3.220	3.329	3.504	3.652	3.845	5.494	5.790	6.049	6.364	6.578	7.012	7.198
	2.907	*2.989*	*3.211*	*3.325*	*3.499*	*3.655*	*3.845*	*5.494*	*5.779*	*6.038*	*6.354*	*6.578*	*7.067*	*7.266*
28	2.952	3.044	3.312	3.424	3.605	3.759	3.953	5.575	5.860	6.113	6.453	6.664	7.138	7.323
	3.020	*3.102*	*3.323*	*3.436*	*3.609*	*3.765*	*3.953*	*5.575*	*5.853*	*6.106*	*6.413*	*6.632*	*7.105*	*7.298*
30	3.084	3.170	3.402	3.511	3.678	3.834	4.017	5.619	5.898	6.170	6.476	6.711	7.205	7.386
	3.088	*3.170*	*3.391*	*3.504*	*3.676*	*3.830*	*4.017*	*5.619*	*5.893*	*6.141*	*6.442*	*6.656*	*7.120*	*7.309*
31	3.149	3.216	3.429	3.541	3.712	3.868	4.055	5.651	5.927	6.180	6.503	6.719	7.146	7.317
	3.127	*3.209*	*3.430*	*3.542*	*3.714*	*3.869*	*4.055*	*5.651*	*5.923*	*6.170*	*6.469*	*6.682*	*7.142*	*7.329*
34	3.216	3.297	3.517	3.630	3.797	3.946	4.127	5.706	5.979	6.224	6.523	6.731	7.218	7.389
	3.202	*3.284*	*3.504*	*3.617*	*3.788*	*3.942*	*4.127*	*5.706*	*5.974*	*6.217*	*6.511*	*6.720*	*7.171*	*7.355*

different choice of quantiles might lead to a better fit for small n; however, a summary of results for the larger values of n is included as Table 4 with the same format as Table 3.

Weibull Distribution Approximation

In fitting a two-parameter Weibull distribution, $F(x) = 1 - e^{-(x/\alpha)\beta}$, to our data, we were precluded from obtaining maximum likelihood estimates since we have at our disposal only the Monte Carlo moments of the distribution and not the individual sample values. In lieu of these, we used a moment estimate suggested by Cohen [3], but with somewhat less than satisfying results. In fact, both the chi approximation and the lognormal seem preferable to the two parameter Weibull, although a three parameter Weibull with the introduction of a location parameter would probably offer some improvement. We felt that this exceeded our charter of consideration of a few commonly encountered distributions, however, and postponed any further inquiry for a later date, especially since the chi and lognormal distributions gave very satisfactory results.

Data Summary and Example of Fitting Procedure

The chi variate is less descriptive in terms of absolute difference between the Monte Carlo and fitted value as we go further out in the tails of the distribution and the parameter n increases. As a matter of fact, the upper tail (perhaps of most interest) is described somewhat better than the lower tail, although the percentage error between the Monte Carloed and fitted value rarely exceeds 4 percent and then only in the extreme percentiles of the lower tail. It is also worthy of note that for hypothesis testing the region of rejection will be slightly larger than that indicated for the fitted χ variate.

The lognormal variate as previously stated provides a good fit for the larger values of $n(15 \leq n \leq 34)$, with a percentage error in excess of 2 percent that occurs only in the most extreme percentiles.

For practical situations either fit is adequate; the chi fit is more versatile over the range of n considered, and the lognormal variate offers a closer approximation over a restricted range of the sample size, n.

Suppose we take a random sample $x_1, \ldots, x_n$ of size n from an *univariate* normal distribution and determine the sample range $w_n = \max_{i,j} |x_i - x_j|$. Since $2m_1/v_1 \cdot w_n^2$ is approximately distributed as $\chi^2(2m_1^2/v_1)$, where $m_1 = E(w_n^2)$ and $v_1 = V(w_n^2)$, we wanted to see if the extreme spread, *ES*, which acts like a bivariate sample range did not also closely follow a chi distribution (or, equivalently, ES^2 follow a χ^2 distribution).

To determine, for example, the 95th percentile $P_{0.95}$ of ES^2 corresponding to some value of n we must satisfy the relation

$$Pr\{ES^2 \leq P_{0.95}\} = 0.95 = Pr\{ES \leq \sqrt{P_{0.95}}\}$$

or equivalently,

$$Pr\left\{\frac{2m}{v} ES^2 \leq \frac{2m}{v} P_{0.95}\right\} = 0.95.$$

To test our approximation, we assume

$$\frac{2m}{v} ES^2 \simeq \chi^2\left(\frac{2m^2}{v}\right),$$

where m and v are the mean and variance of ES^2, so that we may substitute to obtain

$$Pr\left\{\chi^2\left(\frac{2m^2}{v}\right) \leq \frac{2m}{v} P_{0.95}\right\} = 0.95$$

and interpolate in the chi square table to determine $2m/vP_{0.95}$. Finally, multiplication of this quantity by $v/2m$ yields $P_{0.95}$ for ES^2, and $\sqrt{P_{0.95}}$ is the 95-percent point for *ES*.

It is easy to show that the mean, m, and variance, v, for the extreme spread squared (i.e., ES^2) may be expressed in terms of the moments of the extreme spread *ES* as follows:

$$m = \sigma^2 + \mu^2,$$

and

$$v = \alpha_4\sigma^4 + 4\alpha_3\sigma^3\mu + 4\sigma^2\mu^2 - \sigma^4,$$

where the μ, σ, α_3 and α_4 are moments of *ES*. Hence, referring to Table 1 for $n = 10$, we have

$$m = (0.745)^2 + (3.813)^2 = 15.094,$$

$$v = 3.288(0.745)^4 + 4(0.388)(0.745)^3(3.813) + 4(0.745)^2(3.813)^2 - (0.745)^4 = 35.430,$$

and we want to determine $P_{0.95}$, where

$$Pr\{\chi^2(12.86) \leq 0.852P_{0.95}\} = 0.95.$$

nterpolating in the chi square table for 12.86 *d.f.* yields $0.852P_{0.95} = 22.20$, or $P_{0.95} = 26.06$ for ES^2, and $\sqrt{P_{0.95}} = 5.11$ which is the 95-percent point for *ES*, corresponding to the entry in Table 2.

EXAMPLE: Compare the relative precision of the extreme spread *ES* and radial standard deviation *RSD* for 15 rounds.

For $n = 15$ rounds and from Table 1, $\hat{\sigma} = ES/4.190$ gives an unbiased estimate of σ, and the quantity 0.694/4.190 = 0.166 is the relative precision for the extreme spread. In a like manner, the precision of the *RSD* is found from Grubbs [4] Table 4 for 15 rounds to be 0.1817/1.354 = 0.134. Therefore 0.166 vs 0.134 indicates that the *RSD* is slightly more precise than the *ES*. (For the relative precision, we compare standard errors for unbiased estimates.)

To use Table 2, suppose from previous firings we established that $\sigma = 3$ inches. Then for $n = 15$ rounds, the chance that the extreme spread, *ES*, exceeds $5.396\sigma = (5.396)(3) = 16.19$ inches is 0.05.

REFERENCES

1] Aitchison, J. and J. A. C. Brown, *The Lognormal Distribution* (Cambridge University Press, 1966).

2] Cacoullos, T. and H. DeCicco, "On the Distribution of the Bivariate Range," Technometrics *9*(3), 476–80 (1967).

3] Cohen, C. A., "Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples," Technometrics *7*(4), 579–88 (1965).

4] Grubbs. F. E., *Statistical Measures of Accuracy for Riflemen and Missile Engineers* (Edwards Brothers, Ann Arbor, Michigan, 1964). (Copies available from the author.)

5] Grubbs, F. E., H. J. Coon, and E. S. Pearson, "On the Use of Patnaik Type Chi Approximations to the Range in Significance Tests," Biometrika *53*, 248–52 (1966).

A GAME THEORETIC APPROACH TO A TWO FIRM BIDDING PROBLEM*

W. D. Cook

York University
Toronto, Ontario

M. J. L. Kirby

Dalhousie University
Halifax, Nova Scotia

and

S. L. Mehndiratta

Indian Statistical Institute
Madras, India

ABSTRACT

In this paper a model is developed for determining optimal strategies for two competing firms which are about to submit sealed tender bids on K contracts. A contract calls for the winning firm to supply a specific amount of a commodity at the bid price. By the same token, the production of that commodity involves various amounts of N different resources which each firm possesses in limited quantities. It is assumed that the same two firms bid on each contract and that each wants to determine a bidding strategy which will maximize its profits subject to the constraint that the firm must be able to produce the amount of products required to meet the contracts it wins. This bidding model is formulated as a sequence of bimatrix games coupled together by N resource constraints. Since the firms' strategy spaces are intertwined, the usual quadratic programming methods cannot be used to determine equilibrium strategies. In lieu of this a number of theorems are given which partially characterize such strategies. For the single resource problem techniques are developed for determining equilibrium strategies. In the multiple resource problem similar methods yield subequilibrium strategies or strategies that are equilibrium from at least one firm's point of view.

1. INTRODUCTION

This paper deals with a two-firm bidding problem in which each firm has to submit, simultaneously, sealed tender bids for several different contracts. Each contract calls for the winning firm to supply a specific amount of a product at the bid price. It is assumed that the same two firms bid on each contract and that each of the two competing firms wants to determine a bidding strategy which will maximize its profit subject to the constraint that the firm must be able to produce the amounts of the products required to meet the contracts it wins.

Since this bidding problem involves direct competition between the firms, it can be analyzed using game theory notions. This paper develops such a model for determining optimal bidding strategies

*The research on this paper was supported in part under D.R.B. Grant #1200-292-001 and N.R.C. Grant #1200-249-004.

on K contracts, each of which requires the winning firm to produce a commodity whose production involves various amounts of N different resources. The model is formulated as a sequence of bimatrix games coupled together by N constraints, each involving the strategies of both firms. As a result the joint strategy space for the players (firms) is not the cartesian product of their individual strategy spaces. Some results on equilibrium strategies for this bidding game are presented and, in case $N=1$, we show how to compute such strategies. For the multiple resource problem an indication of how to obtain subequilibrium strategies is given.

The present work on the two-firm bidding problem represents an extension in the field of bidding processes to which game theory has been applied. Griesmer and Shubik [5–8], using game theory methods, studied bidding processes which involve submitting bids on a sequence of jobs where there are capacity limitations. The items bid upon are assumed to be indivisible and the capacity limitations are expressed in terms of the number of contracts a firm is capable of supplying. Christenson [1] has used game theory techniques to describe strategic aspects of competitive bidding for corporate securities. For a general reference on game theory approaches to bidding see Rothropf [17] and Stark et al. [19]. Additional references on bidding models are due to Friedman [4], Vickery [20], Wilson [21] and [22] and Kortanek, et al. [9].

2. THE MODEL

Let Q_k denote the (known) amount of the product required to be supplied by the firm that wins the kth contract. Let r_n^k be the amount of resource n, $n=1, 2, \ldots, N$, required by either firm to produce Q_k units of product k. Note that r_n^k may be zero for some n and k. Thus if firm 1 wins the contract to produce product k then firm 1 will be compelled to use up r_n^k units of resource n, $n=1, 2, \ldots, N$, in order to fulfill its contract obligations. In this case firm 2, which has lost the contract, will not have to consume any of its resources on the production of product k.

Assume that the set of admissible bids by firms 1 and 2 on contract k are identical and finite in number. That is, firm 1 considers only bids in the set $\{p_i^k\}_{i=1}^{s_k}$ and firm 2 only bids in the set $\{q_j^k\}_{j=1}^{s_k}$, where $p_i^k=q_i^k$, $i=1, 2, \ldots, s_k$. There is no loss of generality in discretizing the set of bids for a firm, since on a high priced contract, for example, bids will likely be submitted in \$100 or \$1,000 units. On a lower priced contract it is reasonable to assume that bids will be in \$1 or \$10 units. In view of this a firm can estimate quite accurately what the range of bids for the opponent are, and hence can choose the same range of bids for itself. For the sake of convenience, we also assume, with no loss of generality, that p_i^k takes the following form:

$$p_i^k=N^k+i\cdot t^k, \tag{2.1}$$

where N^k and t^k are positive constants. This implies that the difference between any two consecutive admissible bid prices is a constant.

Define the matrix $E^k=(e_{ij}^k)$ by

$$e_{ij}^k=\begin{cases} 1, & i<j \\ 1/2, & i=j \\ 0, & \text{otherwise,} \end{cases} \tag{2.2}$$

where the rows and columns of E^k are indexed by firm 1's and firm 2's admissible bid prices, respectively. Thus E^k represents the win-loss matrix for firm 1, where 1 denotes a win and 0 a loss of the kth contract, and 1/2 indicates that the kth contract splits equally between the firms because of ties in their bid prices.

Using the matrix E^k, we now define another matrix—the total revenue matrix—$P^k=(p_{ij}^k)$ by

$$p_{ij}^k=\begin{cases} p_i^k, & i<j \\ p_i^k/2, & i=j \\ 0, & \text{otherwise,} \end{cases} \tag{2.3}$$

where the entry p_{ij}^k represents the total revenue going to firm 1 from contract k when bids of p_i^k and q_j^k are submitted by firms 1 and 2, respectively. The total revenue matrix for firm 2 is then $(P^k)^T$.*

Consider each pair of matrices $\{P^k, (P^k)^T\}$ as payoff matrices for a bimatrix game, and let X^k and Y^k be mixed strategies for firm (player) 1 and firm (player) 2, respectively. Then the total expected payoff to player 1 is $\sum_k X^k P^k Y^k$ and to player 2 is $\sum_k X^k (P^k)^T Y^k$.

Let

$$R^n=\sum_k r_n^k \qquad \forall n. \tag{2.4}$$

Let M_1^n and M_2^n denote the amounts of resource n available to firms 1 and 2, respectively. Let the resource matrix C_n^k be defined by

$$C_n^k=r_n^k \cdot E^k.$$

The (ij)th component of C_n^k measures the amount of resource consumed by firm 1 in the production of Q_k when bids p_i^k and q_j^k are submitted by firms 1 and 2, respectively.

It is important to stress at this point that the type of bidding environment to which this paper addresses itself is that in which the firms would frequently (say once every month or every quarter) be submitting a series of bids on small, short term contracts. In this situation the process of submitting bids becomes a repeated action, and the basic market structure shows, in general, no drastic change from one senario to the next. It is only in this case that one can utilize the concept of a mixed strategy and, hence, can even consider the game theory approach. Industries in which bids are submitted on large, long term contracts clearly do not fall into this category. Bids would then likely be submitted only once every 2 or 3 years, during which time the entire economic and technological picture of the market could undergo dramatic changes.

The second important issue requiring clarification concerns the meaning of M_1^n and M_2^n. Regardless of whether bidding is done on a month to month basis or once each year, it is rare that a firm would know for certain what its resource limits will be several months prior to the date when these resources

*It should be noted that P^k could just as easily be considered to be a porfit matrix rather than a total revenue matrix, since if c^k is the fixed cost of producing Q_k we could define a matrix $\bar{P}^k=(\bar{p}_{ij}^k)$, where $\bar{p}_i^k=p_i^k-c^k$.

will actually have to be utilized in fulfilling contract commitments. Consequently the M_1^n must be interpreted as "ball park" figures only. Normally, they would represent the *expected* resource availabilities. In other cases firm i may take M_1^n as R percent of what it actually expects to have available with R being a function of the firm's attitude toward risk. Similarily, the requirements r_n^k are only estimates, perhaps on the high side to account for inflation, etc. It is never clear how many dollars of capital, man hours, machine time, . . ., etc. will actually be required or be available to satisfy contract demands.

In view of the above statements, a firm will in almost all cases run either under or over actual limits when the project covered by the contract is complete. In cases of overflow part of the project may have to be subcontracted out. Therefore, the best a firm can hope for is to not exceed resources on "an average basis."

Following the above arguments, we introduce constraints which require that the expected amount of resource n used by each firm not exceed the "amount" of resource available to the firm. Hence the mixed strategy which a firm can use for a given contract is constrained by resource restrictions involving K contracts. These constraints are

$$\sum_k X^k C_n^k Y^k \leqq M_1^n \qquad \forall n$$

and

$$R^n - \sum_k X^k C_n^k Y^k \leqq M_2^n \qquad \forall n,$$

or equivalently*

$$R^n - M_2^n \leqq \sum_k X^k C_n^k Y^k \leqq M_1^n \qquad \forall n. \tag{2.5}$$

Letting $X^k = (x_i^k)$ and $Y^k = (y_j^k)$, we require in addition that

$$\sum_i x_i^k = 1, \qquad x_i^k \geqq 0 \qquad \forall i, k \tag{2.6}$$

$$\sum_j y_j^k = 1, \qquad y_j^k \geqq 0 \qquad \forall j, k. \tag{2.7}$$

---

*It is worth noting that in case lower bounds $\{\underline{M}_1^n, \underline{M}_2^n\}_{n=1}^N$ are imposed on the amounts of resource the firms are allowed to expend the resource constraints (2.5) must be replaced by

$$\hat{M}_2^n \leqq \sum_k X^k C_n^k Y^k \leqq \hat{M}_1^n \qquad \forall n,$$

where $\hat{M}_1^n = \min\{\underline{M}_1^n, R^n - \underline{M}_2^n\}$ and $\hat{M}_2^n = \max\{\underline{M}_1^n, R^n - M_2^n\}$. Thus, lower bounds don't change the general nature of the problem, but merely reduce the size of the feasible region.

Note that (2.5) involves both $\{X^k\}_{k=1}^K$ and $\{Y^k\}_{k=1}^K$. Due to this, the strategy spaces for the players are not independent of each other. Thus the feasibility of a given sequence $\{X^k\}_{k=1}^K$ for firm 1 depends upon the strategy sequence $\{Y^k\}_{k=1}^K$ used by firm 2, and conversely.

In summary, the two firm bidding problem can be viewed as a problem involving a set of K bimatrix games coupled together by N resource constraints (2.5).

3. DEFINITIONS

Let us introduce the following definitions concerning bimatrix games with intertwined strategy spaces.

Definition 1: The strategy sequences $\{X^k\}_{k=1}^K$, $\{Y^k\}_{k=1}^K$ for players 1 and 2, respectively, are *jointly feasible* if they satisfy (2.5), (2.6), and (2.7).

Definition 2: A strategy sequence $\{X_0^k\}_{k=1}^K$ for player 1 is said to be Y-feasible if there exists a sequence $\{Y_0^k\}_{k=1}^K$ such that $\{X_0^k\}_{k=1}^K$, $\{Y_0^k\}_{k=1}^K$ are jointly feasible.

Definition 3: A strategy sequence $\{\hat{Y}^k\}$ for player 2 is said to be X-feasible if there exists a sequence $\{\hat{X}^k\}_{k=1}^K$ such that $\{\hat{X}^k\}_{k=1}^K$, $\{\hat{Y}^k\}_{k=1}^K$ are jointly feasible.

We now define the concept of an equilibrium strategy and a subequilibrium strategy for a sequence of bimatrix games.

Definition 4: Let $\{\hat{X}^k\}_{k=1}^K$, $\{\hat{Y}^k\}_{k=1}^K$ be jointly feasible sequences of strategies for players 1 and 2, respectively. Such sequences are said to constitute an *equilibrium strategy* for the bimatrix games $\{P^k, (P^k)^T\}_{k=1}^K$ if

$$\sum_k \hat{X}^k P^k \hat{Y}^k \geqq \sum_k X^k P^k \hat{Y}^k \tag{3.1}$$

and

$$\sum_k \hat{X}^k (P^k)^T \hat{Y}^k \geqq \sum_k \hat{X}^k (P^k)^T Y^k \tag{3.2}$$

for all jointly feasible sequences $\{X^k\}_{k=1}^K$ and $\{Y^k\}_{k=1}^K$. If only (3.1) or (3.2) is satisfied then $\{X^k\}_{k=1}^K$ and $\{\hat{Y}\}_{k=1}^K$ constitute a *subequilibrium strategy.*

REMARK 1: It may be noted that an equilibrium strategy has the property that a unilateral departure from this strategy by either player will result in a lower expected payoff.

Throughout the remainder of this paper we investigate the problem of finding equilibrium and subequilibrium strategies. The main difficulty arises, not because the sequence of games is constrained, but because the strategy spaces for the players are intertwined. As a result existing techniques for handling bimatrix games [11–14] cannot be applied here. Thus, it is necessary to investigate some of the properties of a sequence of bimatrix games with intertwined strategy spaces. This is done in the following section.

4. PROPERTIES OF BIMATRIX GAMES WITH INTERTWINED STRATEGY SPACES

We begin by proving the following theorem concerning the form of equilibrium points.

THEOREM 4.1: If $R^n \leqq M_1^n + M_2^n$, $n = 1, 2, \ldots, N$ then any equilibrium strategy $\{\hat{X}^k\}_{k=1}^K$, $\{\hat{Y}^k\}_{k=1}^K$, if it exists, for the sequence of bimatrix games $\{P^k, (P^k)^T\}_{k=1}^K$ is such that each $\hat{X}^k$ and $\hat{Y}^k$ has at most two nonzero components and these are adjacent. More precisely, there exist $\{i_k\}_{k=1}^K$, $i_k \in \{1, 2, \ldots, s_k\}$ such that the only nonzero components of $\hat{Y}^k$ are $(\hat{y}^k_{i_k-1}, \hat{y}^k_{i_k})$ and the only nonzero components of $\hat{X}^k$ are either $(\hat{x}^k_{i_k-2}, \hat{x}^k_{i_k-1})$ or $(\hat{x}^k_{i_k-1}, \hat{x}^k_{i_k})$ or $(\hat{x}^k_{i_k}, \hat{x}^k_{i_k+1})$.

PROOF: First assume $K = N = 1$ and let $(\hat{X}, \hat{Y})$ be an equilibrium strategy for the bimatrix game $\{P, P^T\}$. Then

$$\hat{X}P\hat{Y} \geqq XP\hat{Y} \quad \text{and} \quad \hat{X}P^T\hat{Y} \geqq \hat{X}P^TY, \tag{4.1}$$

for all jointly feasible X and Y. Scripts are suppressed in this case for the sake of convenience and clarity.

Let $\hat{y}_{i_*}$ and $\hat{y}_{i_0}$ denote the first and last nonzero components, respectively, of $\hat{Y}$ and $\hat{x}_{i_1}$ the last nonzero component of $\hat{X}$. With no loss of generality, assume that not both $\hat{x}_{i_0}$ and $\hat{x}_{i_0-1}$ are nonzero. Otherwise a strategy Y^0, equivalent to $\hat{Y}$ in terms of expected values, could be defined. Consider the mixed strategy $\bar{Y}$ defined by

$$\bar{y}_{i_*} = \hat{y}_{i_*} - \epsilon_2$$

$$\bar{y}_{i_0-1} = \hat{y}_{i_0-1} + \epsilon_1 + \epsilon_2$$

$$\bar{y}_{i_0} = \hat{y}_{i_0} - \epsilon_1$$

$$\bar{y}_i = \hat{y}_i, \text{ for all other } i,$$

where ϵ_1 and ϵ_2 are positive scalars, such that $\bar{Y}$ is feasible. Of course if $i_* \geqq i_0 - 1$, $\hat{Y}$ is already of the desired form. Hence assume that $i_* < i_0$. We shall show that this contradicts (4.1).

If ϵ_1 and ϵ_2 are selected so that the expected resource usage by the firms remains the same i.e., $\hat{X}C\bar{Y} = \hat{X}C\hat{Y}$, then

$$\epsilon_2 = \left[\frac{\hat{x}_{i_*} + 2 \sum_{i=i_*+1}^{i_0-1} \hat{x}_i + \hat{x}_{i_0-1}}{\hat{x}_{i_0-1} + \hat{x}_{i_0}} \right] \cdot \epsilon_1.$$

By letting $p_{i_0} = p_{i_0-1} + t$, the gain in expected revenue by player 2 who uses $\bar{Y}$ instead of $\hat{Y}$ is

$$\epsilon_1 \left(\left[-\frac{p_{i_*} \cdot \hat{x}_{i_*}}{2} - p_{i_*} \sum_{i=i_*+1}^{i_0-1} \hat{x}_i + \frac{p_{i_0-1}\, x_{i_0-1}}{2} + (p_{i_0-1} - p_{i_*})\hat{x}_{i_0} \right] \right.$$

$$\left. + \left[\frac{\hat{x}_{i_*} + 2 \sum_{i=i_*+1}^{i_0-2} \hat{x}_i + x_{i_0-1}}{\hat{x}_{i_0-1} + \hat{x}_{i_0}} \right] \left[\left(p_{i_0-1} - \frac{p_{i_0}}{2} \right) \hat{x}_{i_0} + \frac{p_{i_0-1}\hat{x}_{i_0-1}}{2} \right] \right\} > 0.$$

Since this contradicts (4.1) it follows that $i_* = i_0 - 1$ or $i_* = i_0$. Hence $\hat{Y}$ is either a pure strategy or else $\hat{y}_{i_0-1}$ and $\hat{y}_{i_0}$ are the only nonzero components.

The form of $\hat{X}$ can be established by means of similar arguments.

For the multiple resource, multiple product case if $\{\hat{X}^k\}_{k=1}^K, \{\hat{Y}^k\}_{k=1}^K$ is an equilibrium strategy, then for $k = k_0$ the N coupling constraints are equivalent to the single constraint

$$R - M_2 \leqq \hat{X}^{k_0} E^{k_0} \hat{Y}^{k_0} = M_1,$$

where

$$R - M_2 = \max_{n \epsilon I^0(k_0)} \left\{ \frac{R^n - M_2^n - \sum_{k \neq k_0} \hat{X}^k C_n^k \hat{Y}^k}{r_n^{k_0}} \right\},$$

$$M_1 = \min_{n \epsilon I^0(k_0)} \left\{ \frac{M_1^n - \sum_{k \neq k_0} \hat{X}^k C_n^k \hat{Y}^k}{r_n^{k_0}} \right\},$$

and

$$I^0(k_0) = \{n | r_n^{k_0} \neq 0\}.$$

The ϵ_1, ϵ_2 argument applied to each pair $(\hat{X}^{k_0}, \hat{Y}^{k_0})$ establishes the result.

In what follows χ denotes the set of all Y-feasible strategies for player 1 and $\mathbf{Y}$ the set of all X-feasible strategies for player 2. Defining

$$\chi(Y) = \left\{ X = (X^1, X^2, \ldots, X^k) \mid R_2^n - M_2^n \leqq \sum_k X^k C_n^k Y^k \leqq M_1^n, \forall n, \sum_i x_i^k = 1, x_i^k \geqq 0 \; \forall i, k \right\}$$

and

$$\mathbf{Y}(X) = \left\{ Y = (Y^1, Y^2, \ldots, Y^k) \mid R^n - M_2^n \leqq \sum_k X^k C_n^k Y^k \leqq M_1^n, \forall n, \sum_j y_j^k = 1, y_j^k \geqq 0 \; \forall j, k \right\},$$

we proceed to establish some auxiliary results which will be needed to obtain equilibrium strategies and subequilibrium strategies for the sequence of bimatrix games.

THEOREM 4.2: Any solution $\{\bar{X}^k\}_{k=1}^K, \{\bar{Y}^k\}_{k=1}^K$ to the problem

$$\min_{Y \epsilon \mathbf{Y}} \max_{X \epsilon \chi(Y)} \sum_k X^k P^k Y^k, \qquad \text{where} \qquad R^n \leqq M_1^n + M_2^n,$$

has the property that at most the first two components of each of the vectors $\bar{X}^k$ and $\bar{Y}^k \; \forall k$ are nonzero.

PROOF: First assume $K = N = 1$ and let $\bar{Y}$ have at most the first two components nonzero. We shall show that any other form of $\bar{Y}$ results in either a lower payoff or the same payoff to player 2.

Let $\bar{X}$ be a strategy for player 1 with the property that $\bar{X}P\bar{Y} \geqq XP\bar{Y} \forall X \epsilon \chi(\bar{Y})$. Assuming that $\bar{y}_2 > 0$, define a new strategy for player 2 by

$$Y^0 = (y_1 - \epsilon_1, y_2 - \epsilon_2, \epsilon_3, \epsilon_4, \ldots, \epsilon_s)^T,$$

where

$$\epsilon_1 + \epsilon_2 = \sum_{i=3}^{s} \epsilon_i.$$

First we consider the case $\epsilon_1 = 0$ and show that there exists a strategy X^0 which gives player 1 a higher payoff than $\bar{X}$ against $\bar{Y}$. This will then prove that $\bar{Y}$ is better than Y^0 from player 2's standpoint.

First note that if $\bar{Y}$ is used by player 2, then by using arguments similar to those in Theorem 4.1 it can be shown that the following hold:

(a) Either $(\bar{x}_1, \bar{x}_2)$ or $(\bar{x}_2, \bar{x}_3)$ are the only nonzero components of $\bar{X}$;
(b) Either $\bar{X}C\bar{Y} = M_1$ or $\bar{x}_1 = 1$;
(c) Either $\bar{x}_3 = 0$ or $r\bar{y}_2 > 2M_1$.

We need only consider the case $r\bar{y}_2 \leqq 2M_1$, since if this is violated player 2 can obtain a higher payoff by decreasing y_2. From (c) it will then follow that $\bar{X}$ has the desired property. Now either $\bar{X}C\bar{Y} = M_1$ or $\bar{x}_1 = 1$. If $\bar{x}_1 = 1$ then $1/2\bar{y}_1 + \bar{y}_2 \leqq M_1/r$ and $\bar{X}$ is optimal for both the old strategy $\bar{Y}$ and the new strategy Y^0. The payoffs to player 1 are the same and hence $\bar{Y}$ is just as favorable to player 2 as is Y^0.

If $1/2y_1 + \bar{y}_2 > M_1/r$ then $\bar{X}C\bar{Y} = M_1$. First assume $1/2(\bar{y}_2 + \epsilon_2) \leqq M_1/r$ in which case player 1's problem against Y^0 is

$$\max \left\{ -\left[\left(p_1 - \frac{p_2}{2}\right)\bar{y}_2 + \frac{p_1}{2}\bar{y}_1 - 1/2\, p_2\epsilon_2\right]\right\} x_2 + p_1[1/2\bar{y}_1 + \bar{y}_2]$$

$$\text{s.t. } 1 \geqq x_2 \geqq \frac{\bar{y}_1 + 2\bar{y}_2 - 2M_1/r}{1 - \epsilon_2}, \qquad x_2 \geqq 0.$$

If the coefficient of x_2 is negative set

$$x_2^0 = \frac{y_1 + 2\bar{y}_2 - 2M_1/r}{1 - \epsilon_2}.$$

Otherwise set $x_2^0 = 1$. In either case it is easily shown that $X^0PY^0 > \bar{X}P\bar{Y}$. Hence $\bar{Y}$ is preferable to Y^0 from player 2's standpoint.

Similar arguments hold if $1/2\ (y_2 + \epsilon_2) > M_1/r$.

To complete the proof we show that $\bar{Y}$ is preferable to $Y^0 = (\bar{y}_1 - \epsilon,\ \bar{y}_2 + \epsilon,\ \bar{0})^T$. If the expected resource usage is to remain the same it is necessary that $X^0 = (\bar{x}_1 - \epsilon,\ \bar{x}_2 + \epsilon,\ \bar{0})^T$. However, this results in an increase in the objective function value by $t/2\ \{\epsilon^2 + \epsilon[2 - (\bar{x}_1 + \bar{y}_1)]\}$.

This proves the result for $N = K = 1$.

In the multiple product case the resource constraints reduce to

$$R - M_2 \leqq \bar{X}^{k_0} E^{k_0} \bar{Y}^{k_0} \leqq M_1$$

for each $k_0 \in \{1, 2, \ldots, K\}$. Applying the above argument to each contract it follows that each of the pairs $(\bar{X}^{k_0}, \bar{Y}^{k_0})$ has the desired property.

The proof of the following theorem proceeds in essentially the same fashion as that of Theorem 4.2 and is therefore omitted.

THEOREM 4.3: Any solution $\{\bar{X}^k\}_{k=1}^K, \{\bar{Y}^k\}_{k=1}^K$ to the problem

$$\max_{X \in \chi} \min_{Y \in \Upsilon(X)} \sum_k X^k P^k Y^k, \qquad \text{where} \qquad R^n \leqq M_1^n + M_2^n \tag{4.3}$$

has the property that at most the last two components in each of the vectors $\bar{X}^k$ and $\bar{Y}^k$ are nonzero.

The following is an immediate consequence of Theorems 4.2 and 4.3.

THEOREM 4.4: (a) any optimal solution $\{\bar{X}^k\}_{k=1}^K, \{\bar{Y}^k\}_{k=1}^K$ to (4.2) is such that for each k either $\bar{X}^k$ or $\bar{Y}^k$ is a pure strategy, with $\bar{x}_1^k = 1$ or $\bar{y}_1^k = 1$, and the other has at most the first two components nonzero.

(b) Any optimal solution $\{\bar{X}^k\}_{k=1}^K, \{\bar{Y}^k\}_{k=1}^K$ to (4.3) is such that for each k either $\bar{X}^k$ of $\bar{Y}^k$ is a pure strategy, with $\bar{x}_{s_k}^k = 1$ or $\bar{y}_{s_k}^k = 1$, and the other has at most the last two components nonzero.

In the following sections we deal with the problem of finding equilibrium strategies for the sequence of bimatrix games $\{P^k, (P^k)^T\}_{k=1}^K$. We shall do this by employing the theorems proved above and considering the problems given in Theorems 4.2 and 4.3. A number of difficulties arise from the fact that

$$\max_{X \in \chi} \min_{Y \in \Upsilon(X)} \sum_k X^k P^k Y^k > \min_{Y \in \Upsilon} \max_{X \in \chi(Y)} \sum_k X^k P^k Y^k$$

which is caused by the familiar Prisoner's Dilemma phenomenon inherent in this problem. As a result, only subequilibrium strategies can be obtained in many cases.

5. EQUILIBRIUM STRATEGIES IN THE SINGLE RESOURCE PROBLEM

Let us consider the case in which $N = 1$ and assume that $R \leqq M_1 + M_2$. We shall show that the solution $\{\bar{X}^k\}_{k=1}^K, \{\bar{Y}^k\}_{k=1}^K$ to (4.2) is an equilibrium strategy for the sequence of bimatrix games $\{P^k, (P^k)^T\}_{k=1}^K$.

By virtue of Theorem 4.2, the coupling constraint (2.5) reduces to

$$R - 2M_2 \leqq \sum_k r^k x_1^k - \sum_k r^k y_1^k \leqq 2M_1 - R. \tag{5.1}$$

Hence the apparent nonlinearity disappears and for the min-max problem (4.2) the strategy spaces are intertwined via a linear constraint on $\{x_1^k\}_{k=1}^K$ and $\{y_1^k\}_{k=1}^K$.

In order to transform (4.2) into a linear programming problem we first note that the system of constraints

$$\sum_k r^k x_1^k \leqq 2M_1 \tag{5.2}$$

and

$$\sum_k r^k y_1^k \leqq 2M_2 \tag{5.3}$$

are redundant in the presence of (2.5), (2.6), and (2.7). Hence, problem (4.2) is equivalent to

$$\begin{aligned} &\min_{Y} \quad \max_{X} \quad \sum_{k} X^k P^k Y^k \\ &\text{s.t.} \quad Y \epsilon \mathbf{Y} \\ &\quad X \epsilon \chi(Y) \\ &\sum_{k} r^k x_1^k \leqq 2M_1 \\ &\sum_{k} r^k y_1^k \leqq 2M_2. \end{aligned} \tag{5.4}$$

It is clear from this new formulation of the min-max problem that at the optimum either

$$\sum_{k} r^k \bar{y}_1^k = 2M_2, \qquad \text{if } 2M_2 \leqq R \tag{5.5}$$

or

$$\bar{y}_1^k = 1 \qquad \forall k. \tag{5.6}$$

If (5.5) holds then from (5.4) $\bar{x}_1^k = 1$ for all k. If (5.6) holds then either $\bar{x}_1^k = 1$ for all k or else $\sum_k r^k \bar{x}^k = 2M_1$ if $2M_1 \leqq R$.

The following theorem summarizes these results.

THEOREM 5.1: The optimal solution $\{\bar{X}^k\}_{k=1}^K$, $\{\bar{Y}^k\}_{k=1}^K$ of (5.4) is such that either

$$\bar{x}_1^k = 1 \ \forall k \qquad \text{and} \qquad \sum_k r^k \bar{y}_1^k = \min\{R, 2M_2\} \tag{5.7}$$

or

$$\bar{y}_1^k = 1 \ \forall k \qquad \text{and} \qquad \sum_k r^k \bar{x}_1^k = \min\{R, 2M_1\}. \tag{5.8}$$

From Theorem 5.1 it follows that (5.4) is equivalent to the linear programming problem

$$\begin{aligned} &\max \sum_{k} (1/2 p_1^k y_1^k - p_1^k) \\ &\text{s.t.} \ \sum_k r^k y_1^k \leqq 2M_2 \\ &\quad 0 \leqq y_1^k \leqq 1 \qquad \forall k. \end{aligned} \tag{5.9}$$

Using (5.9) we shall determine the explicit form of $\bar{y}_1^k$, for all k.

If $R \leqq 2M_2$ and $r^k > 0 \ \forall k$ (the latter can be assumed with no loss of generality), then $\bar{y}_1^k = 1 \ \forall k$. Otherwise $\sum_k r^k \bar{y}_1^k = 2M_2$ and ordering $\left(\frac{p_1^k}{r^k}\right)$ according to

5.10) $$\frac{p_1^{k_1}}{r^{k_1}} \geqq \frac{p_1^{k_2}}{r^{k_2}} \geqq \cdots \geqq \frac{p_1^{k_K}}{r^{k_K}},$$

and defining i_0 to be that index for which

$$\sum_{i=1}^{i_0} r^{k_i} \leqq 2M_2 \qquad \text{and} \qquad \sum_{i=1}^{i_0+1} r^{k_i} > 2M_2,$$

it then follows that

$$\bar{y}_1^{k_i} = \begin{cases} 1 & i = i_0 \\ 2M_2 - \sum_{i=1}^{i_0} r^{k_i}, & i = i_0 + 1 \\ 0, & i > i_0 + 1 \end{cases}$$

and

$$\bar{y}_2^{k_i} = 1 - \bar{y}_1^{k_i}.$$

If (5.8) holds, then (5.4) is equivalent to

(5.11) $$\max \sum_k 1/2 p^k x_1^k$$
$$\text{s.t.} \sum_k r^k x_1^k \leqq 2M$$
$$0 \leqq x_1^k \leqq 1 \qquad \forall k.$$

This problem will give us the explicit form of $\bar{x}_1^k$ and $\bar{x}_2^k$, $\forall k$. It can be easily seen that either $\bar{x}_1^k = 1\ \forall k$, if $R \leqq 2M_1$ or else

$$\bar{x}_1^{k_i} = \begin{cases} 1, & i \leqq i_* \\ 2M_1 - \sum_{i=1}^{i_*} r^{k_i}, & i = i_* + 1 \\ 0, & i > i_* + 1 \end{cases}$$

and

$$\bar{x}_2^{k_i} = 1 - \bar{x}_1^{k_i},$$

where

$$\left\{ \frac{p_1^{k_i}}{r^{k_i}} \right\}_{k=1}^{K}$$

are ordered as in (5.10) and i^* is defined in a manner analogous to i_0.

The proof of the following theorem is a direct consequence of the above arguments.

THEOREM 5.2: The solution $\{\bar{X}^k\}_{k=1}^K$, $\{\bar{Y}^k\}_{k=1}^K$ to (5.4) is an equilibrium strategy for the sequence of bimatrix games $\{P^k, (P^k)^T\}_{k=1}^K$.

One can argue, of course, that the min-max solution does not provide a very satisfactory equilibrium strategy since each player is bidding as low as possible. Of all equilibrium strategies this is likely the

least profitable to either player. One drawback of viewing the bimatrix games in this way is that it assumes player 2 is attempting to be as damaging as possible to player 1. In reality, however, player 2 is interested in maximizing his own gain and is not so much concerned with minimizing his opponent's gain.

A more suitable set of strategies for the players can be obtained by solving the max-min problem (4.3). However, for $R \leqq M_1 + M_2$ it is not necessarily true that the solution to this problem constitutes a true equilibrium strategy. In particular suppose that $R < M_1 + M_2$ and let $\{\bar{X}^k\}_{k=1}^K$ and $\{\bar{Y}^k\}_{k=1}^K$ be an optimal solution to (4.3) for $N=1$. Note that in this case the coupling constraint reduces to

$$R - M_2 \leqq \sum_k r^k y_{s_k}^k - \sum_k r^k x_{s_k}^k \leqq 2M_1 - R, \tag{5.12}$$

and the constraints

$$\sum_k r^k x_{s_k}^k \leqq 2M_2 \tag{5.13}$$

and

$$\sum_k r^k y_{s_k}^k \geqq \min\ \{R,\ 2R - 2M_2\} \tag{5.14}$$

are redundant in the presence of (2.5), (2.6), and (2.7). Hence (4.3) is equivalent to

$$\begin{aligned} &\max_X \min_Y \sum_k X^k P^k Y^k \\ &\text{s.t. } X \epsilon \chi \\ &\quad Y \epsilon \mathbf{Y}(X) \\ &\quad \sum_k r^k x_{s_k}^k \leqq 2M_2 \\ &\quad \sum_k r^k y_{s_k}^k \geqq \min\ \{R,\ 2R - 2M_2\}. \end{aligned} \tag{5.15}$$

If $2M_2 \leqq R$ then at the optimum $\sum_k r^k \bar{x}_{s_k}^k = 2M_2$ and $\bar{y}_{s_k}^k = 1, \forall k$, since $\bar{y}_{s_k}^k = 1 \forall k$ is the only feasible strategy for player 2 in $\mathbf{Y}(\bar{X})$. The reason for this is that as mentioned in Section 4, player 1 wishes to make each $x_{s_k}^k$ as large as possible while player 2 is attempting to make each $y_{s_k}^k$ as small as possible.

With $\bar{y}_{s_k}^k = 1\ \forall k$, $\{\bar{x}_{s_k}^k\}_{k=1}^K$ are found by solving

$$\begin{aligned} &\max \sum_k \left[(1/2p_{s_k}^k - p_{s_{k-1}}^k) x_{s_k}^k + p_{s_{k-1}}^k\right] \\ &\text{s.t. } 0 \leqq x_{s_k}^k \leqq 1 \quad \forall k. \\ &\quad \sum_k r^k x_{s_k}^k = 2M_2. \end{aligned}$$

By ordering

$$\left\{\frac{p^k_{s_k-1}-1/2p^k_{s_k}}{r^k}\right\}^K_{k=1}$$

from smallest to largest it follows that

$$\bar{x}^{k_i}_{s_{k_i}} = \begin{cases} 1, & i \leqq i_0 \\ 2M - \sum_{i=1}^{i_0} r^{k_i}, & i = i_0+1 \\ 0, & i > i_0+1 \end{cases} \tag{5.16}$$

and

$$\bar{x}^{k_i}_{s_{k_{i-1}}} = 1 - \bar{x}^{k_i}_{s_{k_i}} \tag{5.17}$$

$\{\bar{X}^k\}^K_{k=1}$, $\{\bar{Y}^k\}^K_{k=1}$ constitutes a subequilibrium strategy since player 2 cannot deviate from his strategy without violating a resource constraint. However, since player 1 is not utilizing all of his capacity, a unilateral departure from $\{\bar{X}^k\}^K_{k=1}$ to a new strategy which yields a higher expected payoff is possible, and hence this does not constitute an equilibrium strategy.

If $2M_2 > R$ then $\bar{x}^k_{s_k} = 1\ \forall k$ and the objective function of the max-min problem becomes

$$\min \sum_k 1/2\ p^k_{s_k} y^k_{s_k}. \tag{5.18}$$

However, rather than minimizing player 1's payoff, player 2 would be more concerned about maximizing his own payoff and would hence consider $\{(P^k)^T\}^K_{k=1}$ instead of $\{P^k\}^K_{k=1}$. This leads to the linear programming problem

$$\max \sum_k \left[(1/2\ p^k_{s_k} - p^k_{s_{k-1}})y^k_{s_k} + p^k_{s_{k-1}}\right]$$

$$\text{s.t. } 0 \leqq y^k_{s_k} \leqq 1 \qquad \forall k \tag{5.19}$$

$$\sum_k r^k y^k_{s_k} = 2R - 2M_2.$$

The solution to this problem is analogous to (5.16) and (5.17).

Let $\{\bar{Y}^k\}^K_{k=1}$ denote the solution to (5.18) and $\{\hat{Y}^k\}^K_{k=1}$ the solution to (5.19). It follows that $\{\bar{X}^k\}^K_{k=1}$, $\{\hat{Y}^k\}^K_{k=1}$ constitutes a subequilibrium point from player 2's standpoint. It should be noted that since $R < M_1 + M_2$, player 1 is not fully utilizing his capacity and hence could increase his expected payoff by deviating from $\{\bar{X}^k\}^K_{k=1}$. However, assuming that player 2 is playing a feasible strategy, player 1 is assured an expected payoff of at least $\sum_k \bar{X}^k P^k \bar{Y}^k$ and could possibly get $\sum_k \bar{X}^k P^k \hat{Y}^k$. Thus if $2M_2 > R$, $\bar{x}^k_{s_k} = 1\ \forall k$ is a reasonably good strategy for player 1.

In case $R = M_1 + M_2$ then, since all capacity is being fully utilized, the max-min solution $\{\bar{X}^k\}_{k=1}^K$, $\{\bar{Y}^k\}_{k=1}^K$ constitutes a true equilibrium strategy.

Suppose now that the two firms are not capable of jointly absorbing the entire demand. (That is $R > M_1 + M_2$.) In this case the two firms may select some subset K^0 of the products which they do have the capability of supplying, and bid only on those products. It may be that the firms will select the same subset K^0 and a portion of some product $k_0 \notin K^0$ such that their joint capacity is just met. Indeed this is likely to happen since the payoffs are symmetric, and even with no communication between the firms, whatever subset of products is most attractive to one firm will also be most attractive to the other. Thus, at the optimum the firms will select the same set of contracts and portions of contracts to bid on.

Let γ^k denote the fraction of contract k to be bid on. Then

$$0 \leqq \gamma^k \leqq 1 \qquad \forall k. \tag{5.20}$$

We require that

$$\sum_k \gamma^k r^k = M_1 + M_2. \tag{5.21}$$

For the purpose of the following discussion we assume $M_1 \geqq M_2$. In case $M_1 < M_2$ the game can be viewed from player 2's standpoint.

The expected amount of resource used by firm 1 if it randomizes over the last two bids, i.e., $x^k_{s_k-1} + x^k_{s_k} = 1$, and firm 2 bids as high as possible, i.e., $y^k_{s_k} = 1$, is from the definition of C^k,

$$\sum_k (\gamma^k r^k) x^k_{s_k-1} + \sum_k (1/2\ \gamma^k r^k) x^k_{s_k} = M_1. \tag{5.22}$$

The amount used by firm 2 is

$$\sum_k (1/2\ \gamma^k r^k) x^k_{s_k} = M_2.$$

Using (5.23) in (5.22), we obtain

$$\sum_k (\gamma^k r^k) x^k_{s_k-1} = M_1 - M_2.$$

Player 1 now wishes to minimize

$$\sum_k (\gamma^k p^k_{s_k-1}) x^k_{s_k-1} + \sum_k (1/2\ \gamma^k p^k_{s_k}) x^k_{s_k} \tag{5.24}$$

subject to (5.20), (5.21), (5.22), and (5.23) where $\{\gamma^k\}_{k=1}^K$ are treated as variables under player 1's control. This nonlinear programming problem can be converted to a simple linear programming problem if we make the transformations,

(5.25)
$$\begin{cases} \omega^k_{s_{k-1}} = \gamma^k x^k_{s_{k-1}} \\ \omega^k_{s_k} = \gamma^k x^k_{s_k}. \end{cases}$$

Then problem (5.24) is equivalent to the linear programming problem

$$\max \sum_k \left(p^k_{s_{k-1}} \omega^k_{s_{k-1}} + 1/2 p^k_{s_k} \omega^k_{s_k}\right)$$

(5.26)
$$\text{s.t. } \omega^k_{s_{k-1}} + \omega^k_{s_k} \leqq 1 \qquad \forall k$$

$$\sum_k r^k \omega^k_{s_{k-1}} = M_1 - M_2$$

$$\sum_k r^k \omega^k_{s_k} = 2M_2$$

$$\omega^k_{s_{k-1}},\ \omega^k_{s_k} \geqq 0 \qquad \forall k.$$

To see this we note that if $\{\bar{\omega}^k_{s_{k-1}}, \bar{\omega}^k_{s_k}\}_{k=1}^K$ is an optimal solution to (5.26) then

$$\bar{\gamma}_k = \bar{\omega}^k_{s_{k-1}} + \bar{\omega}^k_{s_k},\ \bar{x}^k_{s_{k-1}} = \frac{\bar{\omega}^k_{s_{k-1}}}{\bar{\omega}^k_{s_{k-1}} + \bar{\omega}^k_{s_k}},\ \bar{x}_{s_k} = \frac{\bar{\omega}^k_{s_k}}{\bar{\omega}^k_{s_{k-1}} + \bar{\omega}^k_{s_k}}$$

for $k \in K^1 = \{k \mid \bar{\omega}^k_{s_{k-1}} + \bar{\omega}^k_{s_{k-1}} \neq 0\}$, $\bar{\gamma}^k = 0$, $k \notin K^1$ is an optimal solution to the nonlinear problem (5.20)–(5.24).

THEOREM 5.3: If $\{\bar{\gamma}^k, \bar{x}^k_{s_{k-1}}, \bar{x}^k_{s_k}\}_{k=1}^K$ solves (5.20)–(5.24), then there exists a subset $K^0 \subseteq K^1$ and $k_0 \in K^1 \cap \tilde{K}^0$, where $\tilde{K}^0$ is the complement of K^0, such that $\{\hat{\gamma}^k, \bar{x}^k_{s_{k-1}}, \bar{x}^k_{s_k}\}_{k=1}^K$ is also an optimal solution, where

$$\hat{\gamma}^k = 1, \qquad k \in K^0$$

$$0 \leqq \hat{\gamma}^{k_0} \leqq 1$$

$$\hat{\gamma}^k = 0 \qquad \text{all other } k.$$

PROOF: Assume there exists $k_1, k_2 \in K^1$ such that $\bar{\gamma}^{k_1}, \bar{\gamma}^{k_2}$ are both less than one. If an amount ϵ_1 is subtracted from $\bar{\gamma}^{k_1}$ and ϵ_2 is added to $\bar{\gamma}^{k_2}$ then in order for the total resource usage to remain at $M_1 + M_2$ it follows that $\epsilon_2 = r^{k_1}/r^{k_2}\,\epsilon_1$. The extra payoff received by player 1, if such a change is made, is given by

$$\epsilon_1 \left\{ \frac{r^{k_1}}{r^{k_2}} \left[p^{k_2}_{s_{k_2}-1}\, \bar{x}^{k_2}_{s_{k_2}-1} + 1/2\, p^{k_2}_{s_{k_2}}\, \bar{x}^{k_2}_{s_{k_2}}\right] - \left[p^{k_1}_{s_{k_1}-1}\, \bar{x}^{k_1}_{s_{k_1}-1} + 1/2\, p^{k_1}_{s_{k_1}}\, \bar{x}^{k_1}_{s_{k_1}}\right] \right\}.$$

By the optimality of $\{\bar{\gamma}^k\}_{k=1}^K$ it follows that either k_1 and k_2 do not exist or else the above term in curly

brackets is zero. Hence either $\bar{\gamma}^k = 1\ \forall k \in K^1$ with the exception of possibly one product k_0 or, in case the term in brackets is zero, $\bar{\gamma}^{k_1}$, $\bar{\gamma}^{k_2}$ can be replaced by $\bar{\bar{\gamma}}^{k_1} = \bar{\gamma}^{k_1} - \epsilon_1$ and $\bar{\bar{\gamma}}^{k_2} = \bar{\gamma}^{k_2} + \epsilon_2$, respectively, one of which is equal to 1.

This proves the theorem.

Theorem 5.3 essentially proves that at the optimum the players will select some subset $K^0 \cup \{k_0\} \subseteq \{1, 2, \ldots, k\}$ of contracts and bid only on these, supplying the entire demand for those in K^0 and just enough of k_0 to meet total capacity.

If it is the policy of the buyer that no partial contracts shall be awarded, then $\gamma^k = 0$ or $1, \forall k$ and player 1's problem is equivalent to the mixed integer programming problem

$$\max \sum_k (p^k_{s_k-1}\omega^k_{s_k-1} + 1/2 p^k_{s_k}\omega^k_{s_k})$$

$$\text{s.t. } \omega^k_{s_k-1} + \omega^k_{s_k} = \delta^k \qquad \forall k$$

$$\sum_k r^k \omega^k_{s_k-1} \leqq M_1 - M_2 \tag{5.27}$$

$$\sum_k r^k \omega^k_{s_k} \leqq 2M_2$$

$$\omega^k_{s_k-1}, \omega^k_{s_k} \geqq 0, \qquad \delta^k = 0 \text{ or } 1 \qquad \forall k.$$

Since no partial contracts are allowed it may be that $\{k | \hat{\delta}^k = 1\} \neq K^0$, and there will be some un-utilized capacity. As a result the solution $\{\hat{x}^k_{s_k-1}, \hat{x}^k_{s_k}\}^K_{k=1}$ to (5.27) and $\{\hat{y}^k_{s_k} = 1\}^K_{k=1}$ will constitute a subequilibrium strategy and not necessarily an equilibrium strategy.

In this section we have indicated now to determine equilibrium strategies and subequilibrium strategies for the sequence of bimatrix games $\{P^k, (P^k)^T\}^K_{k=1}$ in the case where $N = 1$. We now turn to a discussion of the multiple resource problem.

6. SUBEQUILIBRIUM STRATEGIES IN THE MULTIPLE RESOURCE PROBLEM

In this section we shall indicate how to determine subequilibrium strategies, or what might appropriately be called "near equilibrium strategies" for the sequence of bimatrix games $\{P^k, (P^k)^T\}^K_{k=1}$ in the N-resources case. The problem of determining equilibrium strategies for this case is difficult for reasons mentioned below.

First assume $R^n \leqq M^n_1 + M^n_2, \forall n$. Unlike the single resource problem, neither the max-min nor the min-max solution will necessarily be an equilibrium strategy to the sequence of bimatrix games. In the single resource case the min-max solution is an equilibrium strategy since a player is either playing a pure strategy or else is at his capacity. In the multiple resource problem, however, a player may be at capacity in one resource, but not in the others. In this case it is possible to make a unilateral departure from the min-max strategy and improve the expected payoff. The min-max solution does, nevertheless, constitute a subequilibrium strategy. However, as mentioned previously, such a strategy is highly unsatisfactory since the expected payoff to each player is very low. Such would also be the case for any equilibrium strategy involving only the first two bids on each product k, provided such a strategy exists.

As was the case for the single resource problem, the maximum solution $\{\bar{X}^k\}_{k=1}^K$, $\{\bar{Y}^k\}_{k=1}^K$ is a more attractive alternative. By fixing $\{\bar{X}^k\}_{k=1}^K$ and solving a linear programming problem for the best strategy $\{\hat{Y}^k\}_{k=1}^K$ for player 2, we get a subequilibrium strategy from player 2's point of view. Player 1 is then guaranteed to get at least $\sum_k \bar{X}^k P^k \bar{Y}^k$ and could possibly get $\sum_k \bar{X}^k P^k \hat{Y}^k$. There is no way of knowing, of course, how close $\{\bar{X}^k\}_{k=1}^K$, $\{\bar{Y}^k\}_{k=1}^K$ or $\{\bar{X}^k\}_{k=1}^K$, $\{\hat{Y}^k\}_{k=1}^K$ is to an equilibrium strategy but they do offer reasonable alternatives.

The max-min problem, analogous to (5.15), for the N-resource, K-contract problem is

$$\max_{X\in\chi} \min_{Y\in\Upsilon(X)} \sum_k X^k P^k Y^k$$

$$\text{s.t. } 0 \leqq x^k_{s_k} \leqq 1 \qquad \forall k$$

$$0 \leqq y^k_{s_k} \leqq 1 \qquad \forall k \tag{6.1}$$

$$R^n - 2M^n_2 \leqq \sum_k r^k_n y^k_{s_k} - \sum r^k x^k_{s_k} \leqq 2M^n - R^n \qquad \forall n.$$

In order to solve (6.1) it is desirable to transform it into a programming problem involving only $\{x^k_{s_k}\}_{k=1}^K$. However, the usual procedures for doing this cannot be immediately applied since the feasibility of a given strategy sequence $\{\bar{Y}^k\}_{k=1}^K$ for player 2 is a function of player 1's strategy sequence and conversely. In order to guarantee that the resulting programming problem is equivalent to (6.1) it is necessary to ensure that the feasible region contains only Y-feasible solutions. Otherwise the problem would be unbounded.

In order to get a problem equivalent to (6.1) we first add the system of redundant constraints

$$R^n - 2M^n_1 \leqq \sum_k r^k_n x^k_s \leqq 2M^n_n \qquad \forall n \tag{6.2}$$

It follows that any strategy for player 1 satisfying (6.2) will be Y-feasible. Adding (6.2) to the system (6.1) and employing the usual techniques [15] it follows that the max-min problem is equivalent to

$$\min \sum_k z^k + \sum_n \left(2M^n_1 - R^n + \sum_k r^k_n x^k_{s_k}\right)\xi^n_1 + \sum_n \left(2M^n_1 - R^n - \sum_k r^k_n x^k_{s_k}\right)\xi^n_2 - \sum_k 1/2p^k_{s_k-1}(1-x^k_{s_k})$$

$$\text{s.t. } z^k + \sum_n \gamma^k_n \xi^n_1 - \sum_n \gamma^k_n \xi^n_2 + \left(\frac{p^k_{s_k} - p^k_{s_k-1}}{2}\right) x^k_{s_k} \geqq -1/2p^k_{s_k-1} \qquad \forall k \tag{6.3}$$

$$0 \leqq x_{s_k} \leqq 1 \qquad \forall k$$

$$R^n - 2M^n_1 \leqq \sum_k r^k_n x^k_{s_k} \leqq 2M^n_2 \qquad \forall n$$

$$z^k \geqq 0 \qquad \forall k; \qquad \xi^n_1, \xi^n_2 \geqq 0 \qquad \forall n.$$

Due to the intertwined nature of the strategy spaces of (6.1), the equivalent problem (6.3) has a nonlinear objective and linear constraints. Although this problem is not as easy to solve as the linear programming problem of the previous section, certain nonlinear programming algorithms, [3] and [16], can be employed to solve it. The optimal solution $\{\bar{X}^k\}_{k=1}^{\kappa}$ to (6.3) will constitute an optimal mixed strategy for player 1 in problem (6.1).

In case $R^n > M_1^n + M_2^n$ for $n \epsilon N' \subseteq \{1, 2, \ldots, N\}$ results similar to those given previously for the single resource problem can be established. However, this case is much more difficult to deal with and, consequently, it shall not be discussed here.

BIBLIOGRAPHY

[1] Christenson, C., "Strategic Aspects of Competitive Bidding for Corporate Securities," Division of Research, Harvard University School of Business, Boston (1965).

[2] Cook, W. D., "Competitive Decision Making Under Uncertainty," Ph.D. Thesis, Dalhousie University, Halifax, Canada (1972).

[3] Everette, H., III, "Generalized Lagrange Multiplier Method for Solving Problems of Optimal Allocation of Resources," Operations Res. *11*, 399–417 (1963).

[4] Friedman, L. A., "A Competitive Bidding Strategy," Operations Res. *4*, 104–112 (1956).

[5] Griesmer, J. H. and M. Shubik, "Toward a Study of Bidding Processes, Part IV: Games with Unknown Costs," Nav. Res. Log. Quart. *14*, 415–433 (1967).

[6] Griesmer, J. H. and M. Shubik, "Toward a Study of Bidding Processes, Part III: Some Special Models," Nav. Res. Log. Quart. *10*, 199–217 (1963).

[7] Griesmer, J. H. and M. Shubik, "Toward a Study of Bidding Processes, Part II: Games with Capacity Limitations," Nav. Res. Log. Quart. *10*, 151–173 (1963).

[8] Griesmer, J. H. and M. Shubik, "Toward a Study of Bidding Processes: Some Constant Sum Games," Nav. Res. Log. Quart. *10*, 11–21 (1963).

[9] Kortanek, K., J. V. Soden, and D. Sodaro, "On a Class of Sequential Probabilistic Bid Pricing Models," Technical Report #66, Dept. of Operations Research, College of Engineering, Cornell University, Ithaca, New York (1969).

[10] Lemke, C. E., "Bimatrix Equilibrium Points and Mathematical Programming," Man. Sci. *11*, 681–689 (1965).

[11] Lemke, C. E. and J. T. Housen, Jr., "Equilibrium Points of Bimatrix Games," SIAM *12*, 413–423 (1964).

[12] Mangasarian, O. L., "Equilibrium Points of Bimatrix Games," SIAM *12*, 778–781 (1964).

[13] Nash, J., "Non-Cooperative Games," Ann. of Math *54*, 286–295 (1951).

[14] Nash, J. F., "Equilibrium Points in *N* Person Games," Proc. Nat. Acad. Sci. U.S.A. *36*, 48–49 (1950).

[15] Owen, G., *Game Theory* (W. B. Saunders, Philadelphia, 1968).

[16] Penn, A. I., "A Generalized Lagrange Multiplier Method for Constrained Matrix Games," Operation Res. *19*, 933–945 (1971).

[17] Rothkopf, M. H., "A Model of Rational Competitive Bidding," Man. Sci. *15*, 362–373 (1969).

[18] Stark, R. M., "Competitive Bidding: A Comprehensive Bibliography," Operations Res. *19*, 484–490 (1971).

[19] Stark, R. M. and R. H. Mayer, Jr., "Some Multicontract Decision Theoretic Competitive Bidding Models," Operations Res. *19*, 469–483 (1971).

[20] Vickery, W., "Counterspeculation, Auctions and Competitive Sealed Tenders," J. Finance *13*, 8–37 (1961).

[21] Wilson, R. B., "Competitive Bidding with Asymmetric Information," Man. Sci. *13*, 816–820 (1967).

[22] Wilson, R. B., "Competitive Bidding with Disparate Information," Man. Sci. *15*, 446–448 (1969).

MULTILINEAR EXTENSIONS AND THE BANZHAF VALUE*

Guillermo Owen

Rice University

ABSTRACT

It is shown that the Banzhaf value can be obtained by differentiating the multilinear extension of a game at the midpoint of the unit cube. This gives us a composition theorem for the value of compound games. As an example, the values of the electoral college and presidential election "games" are approximated by the method of extensions.

MULTILINEAR EXTENSIONS AND THE BANZHAF VALUE

In [5] a multilinear extension was defined for n-person games, as follows: if v is the characteristic function of a game, with player set $N = \{1, 2, \ldots, n\}$, then the multilinear extension is the function of n variables

$$f(x_1, \ldots, x_n) = \sum_{S \subset N} \left\{ \prod_{i \in S} x_i \prod_{i \notin S} (1 - x_i) \right\} v(S) \tag{1}$$

for all $0 \leq x_i \leq 1$, $i = 1, \ldots, n$. It was shown in [5] how this extension is related to the Shapley value, which is obtained as the integral

$$z_i = \int_0^1 f_i(t, t, \ldots, t)\,dt \tag{2}$$

of the partial derivatives of f along the diagonal $x_1 = x_2 = \ldots = x_n$ of the unit cube. This method was used in [6] to approximate the value of a presidential election game.

A different value was defined, at least for simple games by Banzhaf in [2]. Given a simple game with player set N, Banzhaf considers all 2^n possible divisions of the players into two complementary sets, S and $N - S$ (the "yeas" and "nays" in a given vote). For any such division, a player i is said to be *marginal* if, by changing from one set to the other (from S to $N - S$ or vice-versa) he can change S from a winning to a losing coalition or vice-versa. Let, now, η_i denote the number of such partitions $(S, N - S)$ for which i is marginal. Then

$$\beta_i = \frac{\eta_i}{\sum_{j=1}^{n} \eta_j} \tag{3}$$

is the *Banzhaf index* or *voting value* for player i in this game.

*Research supported by Army Research Office – Durham under grant number DA-ARO-D-31-124-72-G105.

A generalization of this value to nonsimple games is quite straightforward. In fact, we see that i is marginal in either of two cases:

(a) $i \in S$, S is winning, $S-\{i\}$ is losing
(b) $i \notin S'$, S' is losing, $S' \cup \{i\}$ is winning.

It is easily seen (by setting $S' = S-\{i\}$ above) that these two cases reduce to the same one. Thus it would be possible to consider only partitions $(S, N-S)$, such that $i \notin S$; then case (a) above could never hold, and we would look only for case (b). This would leave out exactly one half of the marginal cases for player i, but since the value is normalized by (3) there would be no change in the value β_i. Modified in this manner, we see that we have

$$\eta_i = \sum_{\substack{S \\ i \notin S}} [v(S \cup \{i\}) - v(S)], \tag{4}$$

where it is of course understood that $v(S) = 0$ for losing sets, $v(S) = 1$ for winning sets. Since (4) defines the vector η entirely in terms of the characteristic function v, we have a direct generalization to all games in characteristic function form.

As mentioned above, it is only the ratios of the components η_i which matter; thus, we can multiply them by an arbitrary constant, to obtain a new vector ψ:

$$\psi_i = \frac{1}{2^{n-1}} \eta_i$$

or, equivalently,

$$\psi_i = \frac{1}{2^{n-1}} \sum_{\substack{S \\ i \notin S}} [v(S \cup \{i\}) - v(S)]. \tag{5}$$

The advantage of this representation is that there are precisely 2^{n-1} possible coalitions S. Thus the coefficient 2^{1-n} may be thought of as a probability, and we see that ψ_i is the mathematical expectation of the marginal value $v(S \cup \{i\}) - v(S)$, assuming that each of the 2^{n-1} coalitions $S (i \notin S)$ has the same probability (this is in keeping with the usual probabilistic description given for the Shapley value, which of course gives different probabilities to coalitions of different size.)

Let us now differentiate the multilinear extension (1). Letting f_i be the derivative with respect to the ith variable, we have

$$f_i(x_1, \ldots, x_n) = \sum_{\substack{S \subset N \\ i \notin S}} \left\{ \prod_{j \in S} x_j \prod_{\substack{j \notin S \\ j \neq i}} (1 - x_j) \right\} [v(S \cup \{i\}) - v(S)].$$

Setting $x_1 = x_2 = \ldots = x_n = 1/2$, we now have

$$f_i(1/2, 1/2, \ldots, 1/2) = \sum_{\substack{S \subset N \\ i \notin S}} (1/2)^{n-1} [v(S \cup \{i\}) - v(S)], \tag{6}$$

which is precisely ψ_i. Thus we see that, just as the Shapley value is obtained by integrating the gradient of f along the main diagonal of the cube, so the Banzhaf value is obtained by evaluating the gradient at the midpoint of the cube.

One interesting consequence of this result is that, just as the Shapley value can be approximated by the method of multilinear extensions, so also can the Banzhaf value. In fact, it is much easier to approximate the Banzhaf value, as it requires evaluation of the partial derivatives at a single point, whereas the Shapley value requires evaluation along the entire diagonal, plus an integration.

Example: The Electoral College

As an example, consider the "electoral college game," a 51-player weighted majority game in which each player (state) has w_i (between three and 45) votes and 270 votes are needed to win. As pointed out in [5] and [6], the partial derivative $f_i(x_i, \ldots, x_n)$ can be approximated as the probability

$$f_i(x_1, \ldots, x_n) \cong \text{Prob}\{269.5 - w_i \leq Y_i \leq 269.5\}, \tag{7}$$

where Y_i is a normal random variable with mean and variance

$$\mu_i = \sum_{j \neq i} x_j w_j \tag{8}$$

$$\sigma_i^2 = \sum_{j \neq i} x_j(1 - x_j) w_j^2, \tag{9}$$

respectively. At the point $(1/2, 1/2, \ldots, 1/2)$, this is

$$\mu_i = 1/2(538 - w_i) \tag{10}$$

$$\sigma_i^2 = 1/4(9{,}942 - w_i^2) \tag{11}$$

since $\Sigma w_i = 538$ and $\Sigma w_i^2 = 9{,}942$ for the 1970 apportionment. Thus we have the approximation

$$\psi_i \cong \Phi\left(\frac{269.5 - \mu_i}{\sigma_i}\right) - \Phi\left(\frac{269.5 - \mu_i - w_i}{\sigma_i}\right), \tag{12}$$

where Φ is the standard cumulative normal distribution function. Substitution of (10) and (11) here will give us

$$\psi_i = \Phi\left(\frac{w_i + 1}{\sqrt{9{,}942 - w_i^2}}\right) + \Phi\left(\frac{w_i - 1}{\sqrt{9{,}942 - w_i^2}}\right) - 1 \tag{13}$$

as the desired approximation.

TABLE I

Electoral votes (w_i)	Banzhaf value (ψ_i)	Banzhaf ratio (β_i)	Shapley value ($\mathcal{I}_i$)
45	0.38694	0.08828	0.08831
41	0.34806	0.07941	0.07973
27	0.22151	0.05054	0.05096
26	0.21291	0.04857	0.04898
25	0.20435	0.04662	0.04700
21	0.17056	0.03891	0.03917
17	0.13736	0.03134	0.03147
14	0.11276	0.02573	0.02577
13	0.10462	0.02387	0.02388
12	0.09648	0.02201	0.02200
11	0.08838	0.02016	0.02013
10	0.08028	0.01832	0.01827
9	0.07221	0.01647	0.01641
8	0.06414	0.01463	0.01456
7	0.05610	0.01280	0.01272
6	0.04806	0.01096	0.01088
5	0.04004	0.009135	0.009053
4	0.03202	0.007305	0.007230
3	0.02402	0.005480	0.005412

In Table I, the value ψ_i is approximated by this method. Also given are the corresponding $\beta_i = \psi_i / \Sigma \psi_j$, as well as the Shapley values, for comparison.

Behavior Under Composition

In [8], Shapley defines a composition operation for simple games. This composition is generalized in [4] to nonsimple games. Essentially, if v is a simple m-person game, while $u_1, u_2, \ldots, u_m$ are simple games for $n_1, n_2, \ldots, n_m$ persons, respectively, then

$$v[u_1, u_2, \ldots, u_n] = v^* \tag{14}$$

is a simple game for $m^* = n_1 + \ldots + n_m$ persons. The persons are labeled (i, j), with $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, and a set is winning in the game v^* if it contains a subset of the form

$$\bigcup_{i \in T} \{i\} x S_i,$$

where T is winning in v, and each S_i is winning in the corresponding u_i. A rather complicated formula, related to (1), is given in [4] to generalize this to nonsimple games.

It is shown in [5] that, if v^* is the composition (14), then the multilinear extension of v^* is the composite function

$$h = f^0(g_1, g_2, \ldots, g_m), \tag{15}$$

where $f, g_1, g_2, \ldots, g_m$ are respectively the multilinear extensions of $v, u_1, u_2, \ldots, u_m$, respectively.

Here, h is a function of m^* variables, y_{ij}, with $i=1, \ldots, m$ and $j=1, \ldots, n_i$.

Let h_{ij} be the partial derivative of h with respect to y_{ij}. We have

$$h_{ij}=\frac{\partial h}{\partial y_{ij}}=\frac{\partial f}{\partial g_i}\frac{\partial g_i}{\partial y_{ij}}$$

or

16) $$h_{ij}=f_i g_{ij}$$

where f_i is the ith partial of f, while g_{ij} is the jth partial of g_i.

For the Banzhaf value, we must evaluate this at the point $x_{11}=\ldots=x_{mn_m}=1/2$. We have then

17) $$h_{ij}(1/2, \ldots, 1/2)=f_i(x_1, \ldots, x_m)g_{ij}(1/2, 1/2, \ldots, 1/2)$$

where the arguments x_i are given by

18) $$x_i=g_i(1/2, 1/2, \ldots, 1/2).$$

In general, of course, $g_i(1/2, \ldots, 1/2)$ is different from 1/2, and so the Banzhaf value does not, in general, compose. It is interesting to see, however, that the value which is allotted to the members of the ith subgame is distributed among them according to the derivatives of g_i, at the point $(1/2, \ldots, 1/2)$. This is precisely the value for the component game u_i. Thus we have

THEOREM 1: In a compound game $v[u_1, \ldots, u_m]$, the value to members of a subgame u_i is proportional to their value in the subgame itself.

In some cases, we will have, for all i

19) $$x_i=g_i(1/2, 1/2, \ldots, 1/2)=1/2.$$

In this case, of course, the value composes.

In this respect, the following lemma is of interest:

LEMMA: Let u be an n-person game such that $u(N)=1$, and let g be the multilinear extension of u. If u is constant sum, then

$$g(1/2, 1/2, \ldots, 1/2)=1/2.$$

Conversely, if u is super-additive, and $g(1/2, 1/2, \ldots, 1/2)=1/2$, then u is constant-sum.

PROOF: By definition (1),

20) $$g(1/2, 1/2, \ldots, 1/2)=(1/2)^n \sum_{S \subset N} u(S).$$

This sum can be split into two sums, one containing all S such that $1 \epsilon S$, and the other containing the remaining coalitions, which are of course their complements.

$$g(1/2, 1/2, \ldots, 1/2) = (1/2)^n \sum_{\substack{S \\ 1 \in S}} [u(S) + u(N-S)].$$

Now, if u is constant-sum, we have

$$u(S) + u(N-S) = 1,$$

for all S, and so

$$g(1/2, 1/2, \ldots, 1/2) = (1/2)^n \sum_{\substack{S \\ 1 \in S}} 1.$$

There are 2^{n-1} terms in the sum, and so $g(1/2, 1/2, \ldots, 1/2) = 1/2$.

Conversely, suppose u is super-additive. Then

$$u(S) + u(1-S) \leq 1$$

for each S, and so

$$(1/2)^n \sum_{\substack{S \\ 1 \in S}} [u(S) + u(N-S)] \leq 1/2$$

with equality holding only if $u(S) + u(N-S) = 1$ for all S, i.e., if S is constant-sum.

In case of composition, the assumption $u_i(N_i) = 1$ is normally made. Thus, for constant-sum game u_i, we find that (17) takes the form

$$h_{ij}(1/2, 1/2, \ldots, 1/2) = f_i(1/2, \ldots, 1/2) g_{ij}(1/2, 1/2, \ldots, 1/2). \tag{21}$$

Thus we obtain

THEOREM 2: Let v^* be the composition $v[u_1, \ldots, u_m]$, where each u_i is a constant-sum game Let ψ_{ij}^* be the Banzhaf value to player (i, j) in v^*; let ψ_i be the value to player i in game v, and le $\mathscr{I}_{ij}$ be the value to player j in the subgame u_i. Then

$$\psi_{ij}^* = \psi_i \mathscr{I}_{ij}. \tag{22}$$

It should be noted that it is the value ψ (given by (5) or (6)) and not the ratio β (given by (3)) whic composes.

Example: The Presidential Election Game

The presidential election "game" can best be defined as the composition

$$v^* = v[u_1, u_2, \ldots, u_{51}]$$

where v is the electoral college game, described above, and each u_i is a simple majority game for n players, n_i being the number of voters in the ith state ($i = 1, \ldots, 51$).

Strictly speaking, u_i is a constant-sum game only if n_i is odd; with an even number of voters, it is possible to obtain two complementary sets, S and $N-S$, both of which lose. This can only happen, however, if both S and $N-S$ have exactly $n_i/2$ players, and of course, the probability of this, when n_i is large, becomes negligible. Thus each u_i is, for all practical purposes, a constant-sum game. In fact, we note that, for any t,

$$g_i(t, t, \ldots, t) = \text{Prob}\left(z_i > \frac{n_i}{2}\right) \tag{23}$$

where z_i is a binomial random variable with parameters n_i and t. This can be approximated (with negligible error) by a normal variable with mean $n_i t$ and variance $n_i t(1-t)$. Thus

$$g_i(t, \ldots, t) = \Phi\left(\frac{n_i(t-1/2)}{\sqrt{n_i t(1-t)}}\right)$$

or

$$g_i(t, \ldots, t) = \Phi\left(\sqrt{n_i}\,\frac{t-1/2}{\sqrt{t(1-t)}}\right). \tag{24}$$

At $t = 1/2$, this gives us

$$g_i(1/2, 1/2, \ldots, 1/2) = \Phi(0) = 1/2, \tag{25}$$

which is the desired result. Thus Theorem 2 is applicable here.

To obtain the partial derivatives g_{ij}, we note that g_i is symmetric in all its arguments. Thus, for a point on the main diagonal, we have

$$g_{ij}(t, t, \ldots, t) = \frac{1}{n_i}\frac{dg_i(t, t, \ldots, t)}{dt}. \tag{26}$$

Differentiating (24), we have

$$\frac{dg_i(t, \ldots, t)}{dt} = \frac{\sqrt{n_i}}{4(t-t^2)^{3/2}} \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-n_i(t-1/2)^2}{2t(1-t)}\right\} \tag{27}$$

and, at $t = 1/2$, this is

$$\frac{dg_i}{dt} = \sqrt{\frac{2n_i}{\pi}}. \tag{28}$$

Thus

$$g_{ij}(1/2, \ldots, 1/2) = \sqrt{\frac{2}{\pi n_i}} \tag{29}$$

for every $i = 1, \ldots, m$, and $j = 1, \ldots, n_i$.

Application of Theorem 2 now gives us

$$\psi^*_{ij} = \sqrt{\frac{2}{\pi n_i}}\,\psi_i\,, \tag{30}$$

and thus we see that, in the compound (election) game, the value to each voter in the ith state is proportional to the value of the ith state in the quotient (electoral college) game, divided by n_i. Generally, of course, n_i is not known, but if it may be assumed that n_i is proportional to the population p_i of the ith state, then

$$\psi^*_{ij} = \frac{k}{\sqrt{p_i}}\,\psi_i\,, \tag{31}$$

where k is a factor of proportionality. As far as the Banzhaf ratio β_{ij} is concerned, of course, this factor k is unimportant.

In Table II, the values ψ^*_{ij} are computed, setting $k = 1$. (This corresponds, ideally, to setting $n_i = 2p_i/\pi$, so that approximately 63.6 percent of the population votes). Also given are the "state values"

$$\hat{\psi}_i = p_i\,\psi_{ij} = \sqrt{p_i \psi_j}$$

which give the total power of all voters within the state, together. Also given are the ratios,

$$\psi^*_{ij}/\psi_{D.C.,j}$$

which show how the strength of a voter in a given state compares with that of a voter in the least advantaged "state," the District of Columbia, and the ratio $\hat{\beta}_i$, which shows the fraction of the total power accruing to voters in the ith state.

COMMENTS

It is of interest to compare the Banzhaf value for this game with the Shapley value, as obtained in [6].

First of all, we notice that, for the 51-player state game, the Banzhaf value gives slightly less power to the large states than does the Shapley value, though still slightly more than they would receive if the value were directly proportional to the electoral votes.

Next, we note that, in the composed game, the Banzhaf value increases this bias by a factor which is exactly proportional to the square root of the population (number of voters). In the Shapley value, this new factor is not quite equal to $\sqrt{m_i}$. Thus the two values are very nearly equal for the large game. It is not, however, a good idea to take the Shapley value for the state game, and then combine it with the Banzhaf procedure ($\sqrt{m_i}$) for composition, as this will take the "worst" of both values (from the point of view of equity), using each where it is most favorable to the large states.

Finally, there may be some question as to the probable error caused by the approximations (12). This is, of course, difficult to evaluate; we point out, however, that in [6] a similar approximation

TABLE II

State	Electoral votes	Population	Value per vote ψ_{ij}^{*} ($\times 10^{-5}$)	$\frac{\psi_{ij}^{*}}{\psi_{D.C.,j}}$	Total to state $\sqrt{m_i}\,\psi_i$	Portion to state β_i
Alabama	9	3,444,165	3.8909	1.409	134.01	0.01192
Alaska	3	302,173	4.3696	1.582	34.01	0.00117
Arizona	6	1,772,482	3.6099	1.307	63.98	0.00569
Arkansas	6	1,923,295	3.4655	1.255	66.65	0.00593
California	45	19,953,134	8.6624	3.137	1,728.42	0.15371
Colorado	7	2,207,259	3.7760	1.367	83.35	0.00741
Connecticut	9	3,032,217	4.1468	1.502	125.74	0.01118
Delaware	3	548,104	3.2445	1.175	17.78	0.00158
District of Columbia	3	765,510	2.7616	1.000	20.89	0.00186
Florida	17	6,789,443	5.2716	1.909	357.91	0.03183
Georgia	12	4,589,575	4.5035	1.631	206.69	0.01838
Hawaii	4	769,913	3.6524	1.323	28.07	0.00250
Idaho	4	713,008	3.7932	1.374	27.03	0.00240
Illinois	26	11,113,976	6.3865	2.313	709.79	0.06312
Indiana	13	5,193,669	4.5907	1.662	238.42	0.02120
Iowa	8	2,825,041	3.8161	1.382	107.81	0.00959
Kansas	7	2,249,071	3.7408	1.355	84.13	0.00748
Kentucky	9	3,219,311	4.0245	1.457	129.56	0.01152
Louisiana	10	3,643,180	4.2060	1.523	153.23	0.01363
Maine	4	993,663	3.2148	1.164	31.89	0.00284
Maryland	10	3,922,399	4.0535	1.468	159.99	0.01414
Massachusetts	14	5,689,170	4.7275	1.712	268.95	0.02392
Michigan	21	8,875,083	5.7252	2.073	508.12	0.04519
Minnesota	10	3,805,069	4.1155	1.490	156.60	0.01393
Mississippi	7	2,216,912	3.7678	1.364	85.53	0.00743
Missouri	12	4,677,399	4.4610	1.615	208.66	0.01856
Montana	4	694,409	3.8407	1.391	26.70	0.00237
Nebraska	5	1,483,791	3.2871	1.190	48.77	0.00434
Nevada	3	488,738	3.4359	1.244	16.79	0.00149
New Hampshire	4	737,681	3.7281	1.350	27.50	0.00245
New Jersey	17	7,168,164	5.1305	1.858	367.76	0.03271
New Mexico	4	1,016,000	3.1767	1.150	32.28	0.00287
New York	41	18,190,740	8.1607	2.955	1,484.50	0.13202
North Carolina	13	5,082,059	4.6408	1.680	235.85	0.02097
North Dakota	3	617,761	3.0561	1.107	18.88	0.00168
Ohio	25	10,652,017	6.2612	2.267	666.95	0.05931
Oklahoma	7	2,559,253	3.5068	1.270	89.75	0.00798
Oregon	6	2,091,385	3.3233	1.203	69.50	0.00618
Pennsylvania	27	11,793,909	6.4501	2.336	760.72	0.06765
Rhode Island	4	949,723	3.2857	1.190	31.20	0.00277
South Carolina	8	2,590,516	3.9851	1.443	103.23	0.00918
South Dakota	4	666,257	3.9250	1.421	26.12	0.00232
Tennessee	10	3,924,164	4.0526	1.467	159.03	0.01414
Texas	26	11,196,730	6.3628	2.304	712.43	0.06336
Utah	4	1,059,273	3.1111	1.127	32.96	0.00293
Vermont	3	444,732	2.6035	1.305	16.01	0.00142
Virginia	12	4,648,494	4.4749	1.620	208.01	0.01850
Washington	9	3,409,169	3.9109	1.416	133.33	0.01186
West Virginia	6	1,744,237	3.6390	1.318	63.47	0.00564
Wisconsin	11	4,417,933	4.2048	1.523	185.76	0.01652
Wyoming	3	332,416	4.1661	1.509	13.85	0.00123

to the Shapley value gave errors which were uniformly less than 0.4 percent. Thus we have reason to believe that the error here should be no greater than one half of one percent in any case.

BIBLIOGRAPHY

[1] Banzhaf, John F., III, "Weighted Voting Doesn't Work: A Mathematical Analysis," Rutgers Law Review, *19*, 317–343 (1965).

[2] Banzhaf, John F., III, "One Man, 3.312 Votes: A Mathematical Analysis of the Electoral College," Villanova Law Review *13*, 304–332 (1968).

[3] Lucas, W. F., "Measuring Power in Weighted Voting Systems," unpublished manuscript, Cornell University (Oct. 1973).

[4] Owen, G., "The Tensor Composition of Non-Negative Games," *Annals of Mathematics Study 52* (Princeton University Press, 1964), pp. 307–326.

[5] Owen, G., "Multi-Linear Extensions of Games," Management Science *18*, P64–P79 (1972).

[6] Owen, G., "Evaluation of a Presidential Election Game," Rice University, July 1974 (submitted for publication.

[7] Shapley, L. S., "A Value for n-Person Games," *Annals of Mathematics Study 28* (Princeton University Press, 1953), pp. 307–317.

[8] Shapley, L. S., "Solutions of Compound Simple Games," *Annals of Mathematics Study 52* (Princeton University Press, 1964), pp. 267–306.

TRANSPORTATION PROBLEMS WITH SOME x_{ij} NEGATIVE AND TRANSSHIPMENT PROBLEMS

Paul S. Dwyer*

The University of Michigan

ABSTRACT

The general solution process of the Hitchcock transportation problem resulting from the application of the method of reduced matrices may give solutions with some negative x_{ij} values. This paper is devoted to a review of the reduced matrices method, an examination of suitable interpretation of sets of x_{ij} which include some negative values, and ways of interpreting these values in useful modifications of the Hitchcock problem. Such modifications include a) the reshipment problem, b) the overshipment problem, and c) the transshipment problem. Techniques are developed for determining and eliminating c_{ij} which are not optimal. These techniques and results are useful in solving the problems indicated above. The natural applicability of the simple and general method of reduced matrices is emphasized.

1. INTRODUCTION

This paper results from a continuation of study on the solution process for the multidimensional Hitchcock transportation problem with the application of the method of reduced matrices [5] and [9]. More specifically, it is a sequel to the paper on the specific application of the method of the two dimensional case [6] which produces general solutions and does not impose, at the beginning of the solution process, the condition $x_{ij} \geq 0$. The present paper is devoted to an examination of the results of this process and to their usefulness in practical problems. One interpretation, that a negative x_{ij} value indicates a (reverse) shipment from destination j to origin i, while not consistent with the Hitchcock condition that parcels can be shipped only from origins to destinations, is in agreement with the transshipment concept that parcels may be shipped more generally.

This paper provides a method for discovering and solving those problems in which a reshipment gives a smaller total cost T_r than the optimal T_h of the Hitchcock solution, problems in which overshipments may lead to cost $T_0 < T_h$, and to transshipment problems generally. Important to the development is the determination of "nonoptimal" c_{ij}, where present, and their replacement by elements which are "barely optimal."

2. REVIEW OF THE METHOD OF REDUCED MATRICES

Since this paper is essentially a sequel to [6], and since some adjustment in the method seems appropriate for solutions which seek suitable negative x_{ij} values, a review of the method, as applied to the Hitchcock problem with $x_{ij}^{(h)} \geq 0$, seems in order. The nature of and objectives of the main steps of and method are outlined. Details and illustrations are available in [6].

According to the Hitchcock problem in which shipments of x_{ij} parcels with costs $-\infty < c_{ij} < \infty$

---

* Professor Emeritus

(though usually $0 < c_{ij} < \infty$) can be made only from origins i, $1 \leq i \leq m$ to destinations j, $1 \leq j \leq n$, it is desired to determine all sets of values $x_{ij}^{(h)} \geq 0$, hereafter called solutions, which satisfy

$$\sum_i x_{ij}^{(h)} = b_j > 0, \qquad \sum_j x_{ij}^{(h)} = a_i > 0, \qquad \sum_i a_i = \sum_j b_j = N \tag{2.1}$$

so that

$$T_h = \sum_{i,j} x_{ij}^{(h)} c_{ij} \tag{2.2}$$

is as small as possible. This statement of the problem differs from a common one in which only one set of $x_{ij}^{(h)} \geq 0$, a particular solution, is required.

The basic approach with the method of reduced matrices is very simple. From the cost matrix, $C = ||c_{ij}||$, one subtracts constants, U_i from row i and V_j from column j, to produce a reduced cost matrix,

$$c_{ij}^{(t)} = c_{ij} - U_i - V_j \tag{2.3}$$

which necessarily has the same feasible and optimal values of $x_{ij}^{(h)}$ as C and also has total transformed cost

$$T_h^{(t)} = \sum_{ij} x_{ij}^{(h)} c_{ij}^{(t)} = T_h - \sum_i a_i U_i - \sum_j b_j V_j. \tag{2.4}$$

More specifically, the method is directed toward the determination of values of U_i and V_j which result in $c_{ij}^{(t)} \geq 0$. By successive subtractions, described in the STEPS below, one eventually arrives at a $C^{(t)}$ matrix which has enough zero elements so that each of the $x_{ij}^{(h)} > 0$ values satisfying (2.1) is associated with an (i, j) element having $c_{ij}^{(t)} = 0$. Then the orthogonality condition

$$x_{ij}^{(h)} c_{ij}^{(t)} = 0 \tag{2.5}$$

is satisfied for all $x_{ij}^{(h)}$. Then by (2.4), $T_h^{(t)} = 0$ and there can be no smaller $T_h^{(t)}$. So

$$T_h = \sum_i a_i U_i + \sum_j b_j V_j$$

is the smallest possible cost. The reduction of the C matrix is complete and the $C^{(t)}$ is said to be "completely reduced."

The solution process thus reduces to the determination of the quantities U_i and V_j which satisfy the conditions above. This is accomplished by a series of simple STEPS. It is obviously necessary that there be at least one zero term in each line (row or column) so STEP I consists in arriving at an "initially reduced matrix" which has the property and, of course, all transformed $c_{ij} \geq 0$. This is accomplished easily in various ways though the result is not necessarily (and need not be) unique. For

example, one may subtract first the smallest value in each row, and then the smallest value in each resultant column, or vice versa, to arrive at the values of $u_i^{(0)}$ and $u_j^{(0)}$. The "reduction" is then measured by the net weighted subtraction

$$R_h^{(0)}=\sum_i a_i u_i^{(0)}+\sum b_j v_j^{(0)}. \tag{2.6}$$

Obviously the initially reduced matrix is not generally unique, since different sets of $u_i^{(0)}$ and $v_j^{(0)}$ can be used and, indeed, the reduction $R_h^{(0)}$ is not unique either. Since T_h is an upper bound for the reduction, it may be preferable to work toward making $R_h^{(0)}$ large. One way of doing this is to select the lines sequentially in the order of their contribution to the reduction, as Mueller-Merbach has suggested as an initial reduction, before application of the Ford-Fulkerson method, [12]. All that is required for STEP I, however, is the determination of some $u_i^{(0)}$, $v_j^{(0)}$ with $c_{ij}^{(1)}=c_{ij}-a_i u_i^{(0)}-b_j v_j^{(0)}\geq 0$.

It is also necessary that there be enough zero terms in the reduced matrix, *for the Hitchcock problem*, so that the positive assignments which may be available for the zero terms of any line be large enough to meet the specifications (a_i, b_j) for that line. STEP II is designated to provide values $u_i^{(1)}$ and $v_j^{(1)}$ so that this condition is satisfied. We let

$$\delta_{ij}=\begin{cases}1 & \text{where } c_{ij}^{(1)}=0\\ 0 & \text{otherwise}\end{cases}$$

and define, for column j

$$g_j=\sum_i a_i\delta_{ij}.$$

It is necessary, for nonnegative $x_{ij}^{(h)}$, that

$$D_j=g_j-b_j\geq 0. \tag{2.7}$$

Where $D_j<0$ for some column, we say there is a deficiency of zero terms in that column. A similar calculation provides D_i, a measure of deficiency for row i. The first phase of STEP II consists in calculating all D_j and D_i. If all D_j and D_i are nonnegative, STEP II is complete. Otherwise, we select the line having the largest $|D_i|a_i$ or $|D_j|b_j$ from those with D_i and/or D_j negative.

Once the line (say column J) is selected, the second phase of STEP II carries out the reduction. The smallest positive element of $c^{(1)}$, denoted by $v_J^{(1)}$ is determined and subtracted from column J. Then $-v_J^{(1)}$ is subtracted from all rows in which $c_{iJ}^{(1)}=0$, to maintain $c_{iJ}^{(2)}\geq 0$. The increment to the reduction is $(b_J-g_J)v_J^{(1)}=D_J v_J^{(1)}>0$. If any other column has resulting elements all positive, the smallest of these is subtracted from that column. If the line selected is I, the treatment is similar with the role of column and row interchanged.

This process is applied sequentially till all D_i and D_j are nonnegative. The resulting matrix $C^{(2)}$ is then said to be "reduced grouped." If $U_i^{(1)}$ and $V_j^{(1)}$ represent the net of all values subtracted in STEP I and STEP II, we have, as the reduction

$$R_h^{(1)} = \sum_i a_i U_i^{(1)} + \sum_j b_j V_j^{(1)}. \tag{2.8}$$

It is to be noted that STEP II is not required for Hitchcock problems with $a_i = b_j = 1$, sometimes called assignment problems, since there is no grouping of individual parcels to form a_i and b_j. Also, in any problem, it is not *necessary* to perform STEP II since one may move at once, upon the completion of STEP I, into the more involved STEP III. However, where Hitchcock solutions with $x_{ij}^{(h)} \geq 0$ are desired, it seems wise to increase the reduction by using the simple transformations of STEP II before undertaking the more involved transformations of STEP III. The transformations of STEP I and STEP II are, in a sense, preliminary transformations which prepare the problem for the effective use of direct substitution in (2.1). Through the values of $U_j^{(1)}$, $V_j^{(1)}$, and $R_h^{(1)}$ are not necessarily unique, it should be pointed out that the application of the first two STEPS of the method of reduced matrices produces reduced grouped matrices which provide Hitchcock solutions for many trivial problems and for other problems which are substantial enough to be used as illustrations for other methods. See Tables 1, 2, and 6 of [6] in which the first two STEPS give general solutions for illustrations of Balinski-Gomory [1], Dantzig [3], and Charnes-Kirby [2].

A presentation of the *formal* method of solution of STEPS III and IV using the direct algebraic solution of (2.1) is available in the earlier papers. In this paper it is shown, less formally, how to make these final reductions using graph theoretic concepts which are especially appropriate to situations in which some of the x_{ij} may be negative.

In the first phase of STEP III, we examine the reduced grouped matrix to see if sets of x_{ij} satisfying (2.1) exist. This is easily done. If every zero term in the $C^{(2)}$ matrix can be connected to every other zero term by lines which join zero terms in the same row (column), then integral x_{ij} (not necessarily positive) can be found. In addition, even though this condition is not satisfied for all zero terms of the matrix, such sets of x_{ij} values can be found when every subset of rows and columns with the property just stated has the sum of the associated a_i equal to the sum of the associated b_j. More formally, if I_k is a complete subset of connected rows and J_k the connected subset of columns, then integral solutions, $x_{I_k J_k}$, though not necessarily positive, exist when $\Theta_{I_k} a_i = \Theta_{J_k} b_j$. This follows from the fact that under the conditions specified, the transforms of Equations (2.1) are solvable for sets of integral x_{ij}.

Using graph theoretic notation, we can say that a node set consists of nodes $\{A_i\}^m$ corresponding to rows of $C^{(2)}$ and nodes $\{B_j\}^n$ corresponding to columns, while the link set consists of those pairs (A_i, B_j) for which $c_{ij}^{(2)} = 0$. Then the graph $G^{(2)}$ has the node set $\{A_i : i \epsilon I\} \cup \{B_j : j \epsilon J\}$ and $G_k^{(2)}$ are the connected components $1 \leq k \leq K$, where K is not greater than $\min(m, n)$. We associate with each $G_k^{(2)}$ the expression $L_k^{(2)} = \Sigma_{I_k} a_i - \Sigma_{J_k} b_j$. When all $L_k^{(2)} = 0$ (which includes the case with $K = 1$) complete sets of integral x_{ij} are computable. When $L_k^{(2)} \neq 0$ for some values of k, $\Sigma_{I_k} a_i \neq \Sigma_{J_k} b_j$ and further reductions are required.

In the first phase of STEP III we compute all values of $L_k^{(2)} \neq 0$, (though alternatively the determination of a single $L_k^{(2)} \neq 0$ is adequate for the sequential process). Then $\Sigma_K L_k^{(2)} \neq 0$ can be used as a check. For some k, say k^* for which $L_{k^*} \neq 0$, we subtract a positive constant from rows I_{k^*} when $\Sigma_{I_{k^*}} a_i > \Sigma_{J_{k^*}} b_j$, or from the columns when $\Sigma_{J_{k^*}} b_j > \Sigma_{I_{k^*}} a_i$. The positive quality subtracted, say Δ, is the smallest nonzero $c_{ij}^{(2)}$ in $I_{k^*} \cap \bar{J}_{k^*}$ (or in $J_{k^*} \cap \bar{I}_{k^*}$), where $\bar{J}_{k^*}$ indicates the complementary set of J_{k^*}. To maintain the $c_{ij}^{(t)} \geq 0$ property we also subtract $-\Delta$ from the columns of J_{k^*} (rows of I_{k^*}). The net positive reduction is $L_{k^*}\Delta$ when L_{k^*} is positive and $-L_{k^*}\Delta$ when L_{k^*} is negative. Hence the application

of this phase of STEP III leads to an additional reduction in that the value of $R_h^{(1)}$ is increased by $\Delta|L_{k^*}|$ and the value of K is decreased by at least one, since at least two of the components are linked. STEP III is applied sequentially until all L_{k^*} are zero and, usually $K=1$. The number of such sequential applications cannot be greater than $\min(m, n)-1$ when the process is applied to the initially reduced matrix. It is usually very much less than this when the process is applied to a matrix which is reduced grouped.

If desired, the reduction may be made in the order indicated by the largest $|L_{k^*}|$, or the largest $|L_{k^*}|\Delta_{k^*}$. This may be advantageous in obtaining large reductions early in STEP III, but it is generally quite satisfactory to work with any $L_{k^*} \neq 0$.

The termination of the first phase of STEP III leads to a matrix $C^{(3)}$ which might be called "consistent reduced" since a solution can be obtained having sets of integral x_{ij} (but not necessarily sets of integral $x_{ij}^{(h)} > 0$) by using all $x_{ij} \neq 0$ associated with $c_{ij}^{(2)} = 0$ terms.

The second phase of STEP III consists in producing these values of x_{ij}. In large problems this may be done by using the reduced matrices method with a computer [8]. Less formally we use the following plan sequentially. If there is a line with a single unassigned zero term, we first assign the a_i or b_j of that line to that term and apply sequentially until all such assignments are made. Then we consider lines with two unassigned zero terms. We select one of these terms corresponding to element $\alpha\beta$, and assign the unspecified $x_{\alpha\beta}$ to it and make additional assignments, which are not necessarily positive, in this manner. We continue until all $c_{ij}=0$ elements have assignments and check, finally, by showing that (2.1) is satisfied.

It should be emphasized that there is no guarantee that the solution, the set of x_{ij}, is unique. In many problems one (or more) of the x_{ij} values may be represented by the general $x_{\alpha\beta}$. Furthermore different solutions with negative elements, some with different total sums T_c, are possible at the end of STEP III.

In many problems the sets of x_{ij} resulting from the solution of the consistent equations are all nonnegative. Then x_{ij} form the general $x_{ij}^{(h)}$ and the Hitchcock solution is complete. In some problems, the unspecified $x_{\alpha\beta}$ in the general solution can take on both positive and negative values. In this case the general Hitchcock solution consists of all sets of $x_{ij}^{(h)} \geq 0$. However, the total cost T_h is unique and is equal to R_h.

When this condition is not satisfied, the specifications of a Hitchcock solution require the elimination of the negative $x_{\alpha\beta}$. This is accomplished in STEP IV which is initiated with the so-called negative solution transformation. We eliminate the negative $x_{\alpha\beta}$ by the transforming $c_{\alpha\beta}^{(3)}$ from zero by subtraction of a constant from row α, or one from column β. To determine which, and the constant to be subtracted, we consider the matrix $C^{(3)}$ with $c_{\alpha\beta}^{(3)} \neq 0$, thus generally increasing the value of K by unity. We then apply the methods of STEP III to determine the transformation eliminating the negative $x_{\alpha\beta}$ and, at the same time, increasing the reduction. In some cases it may be necessary to complete more than one negative solution transformation but the process leads eventually to the general solution, $x_{ij}^{(h)}$ of the Hitchcock problem.

3. MODIFIED PROBLEMS USING NEGATIVE x_{ij}

Instead of eliminating the negative value of x_{ij} to produce the Hitchcock solution, as is done in STEP IV, we here explore the possibility of modifying the problem in such a way that the negative x_{ij} provide a useful result. One such modification is to make the positive assignments and to ignore the

negative ones. The resulting solution might be called an "overshipment" solution since it results in more parcels shipped than required. It should be noted that all the shipping requirements are met (some with a surplus) with a total cost, say T_0, that may be less than T_h of the Hitchcock solution.

There may be many practical problems, for example when periodic shipments are made, in which occasional overshipments may be sensible. Kuhn and Baumol [11] have proposed overshipments of this sort in order to increase the order of degeneracy for fixed charges problems and Dwyer [7] has also featured overshipments in his treatment of such problems with reduced matrices.

The formal determination of the amount of the overshipment is very easy for each known solution with one or more negative x_{ij}. We simply add the absolute value of the negative x_{ij} to the corresponding a_i, and to b_j. For example with $x_{24}=-3$, we add three units to a_2 (which means three additional units are shipped from origin two) and also three units to b_4 (which means that destination four receives three more units than specified).

A formula for T_0 is easily arrived at. If T_c is the (algebraic) cost of the shipment with some negative x_{ij} computed using (2.2) with $x_{ij}^{(h)}$ replaced by x_{ij}, then for $c_{ij} \geqslant 0$

$$T_0 = T_c + \Sigma |x_{ij}| c_{ij} = T_c + \Sigma x_{ij}(-c_{ij}) \tag{3.1}$$

where the summation holds for all i, j with x_{ij} negative. The value of T_0 is not necessarily smaller than T_h, but it sometimes is. The question of determining when $T_0 < T_h$ and the maximum size of $T_h - T_0$ is an important concern of this paper and is treated in a later section.

Since a positive x_{ij} indicates a shipment from origin i to destination j, it seems logical to modify the Hitchcock problem by interpreting the negative x_{ij} as a shipment from destination j to origin i, which can be done after an overshipment to destination j is made. Indeed, the admission of such a reverse shipment permits a formal solution to the modified problem with some x_{ij} negative for the specified a_i, b_j. Again the cost T_c must be recognized as needing interpretation since it contains the term $x_{ij}c_{ij}$ which is negative when $x_{ij} < 0$ and $c_{ij} > 0$. However, the general results are immediately applicable to all situations in which the reverse shipment cost equals the direct shipment cost, $c_{ji} = c_{ij}$, and to other situations in which c_{ji} is known. With the total reshipment cost T_r having been computed by applying a positive cost for each reshipment, we can write the general formula, when $c_{ij} \geqslant 0$

$$T_r = T_0 + \Sigma |x_{ij}| c_{ji} = T_c + \Sigma |x_{ij}| (c_{ij} + c_{ji}) = T_c + \Sigma (x_{ij})\ (-c_{ij} - c_{ji}) \tag{3.2}$$

where the summation holds for all i, j for which $x_{ij} < 0$. If $c_{ji} = c_{ij}$ for $i \neq j$, (3.2) becomes

$$T_r = T_0 + \Sigma |x_{ij}| c_{ij} = T_0 + \Sigma x_{ij}(-c_{ij}) = T_c + 2\Sigma |x_{ij}| c_{ij} = T_c + 2\Sigma x_{ij}(-c_{ij}). \tag{3.3}$$

Is it possible that the reverse shipment solution may have smaller total cost than the accepted Hitchcock solution, commonly considered to be optimal? The answer, for some problems, is "yes," and depends on the value of the c_{ij} associated with the negative x_{ij} and the value of $\Delta_T = T_h - T_c$, the additional reduction necessary to reduce the consistent solution to the Hitchcock solution. We have, with $c_{ij} \geqslant 0$, $c_{ji} \geqslant 0$

$$\left\{\begin{array}{ll} T_r < T_h & \text{when } \Sigma (x_{ij})\ (-c_{ij} - c_{ji}) < \Delta_T \\ T_0 < T_h & \text{when } \Sigma (x_{ij})\ (-c_{ji}) < \Delta_T \end{array}\right\} \tag{3.4}$$

A simple illustration with both T_r and T_0 appreciably less than T_h is presented in Table 1. The initial transformation with $u_1^{(0)}=1$, $u_2^{(0)}=3$, $v_2^{(0)}=1$ results in the consistent reduced matrix $C^{(3)}$. The consistent solution, shown by the right superscripts of the $C^{(3)}$ matrix, has $x_{11}=-7$. Removal of this by the techniques of STEP IV, specifically the subtraction of -4 from row 1 with associated adjustment (subtraction of 4 from column 2) leads to the completely reduced matrix $C^{(4)}$ and the Hitchcock solution indicated.

TABLE 1. *A problem with* T_0 *and* $T_r < T_h$.

C

a_i \ b_j	2	8	$u_i^{(0)}$	$U_i^{(0)}$
1	1	2	1	1
9	3	8	3	3
$v_j^{(0)}$		1	36	
$V_j^{(0)}$		1		36

$C^{(3)}$

a_i \ b_j	2	8	$u_i^{(3)}$	$U_i^{(3)}$
1	0^{-7}	0^{8}	-4	-3
9	0^{9}	4		3
$v_j^{(3)}$		4	28	
$V_j^{(3)}$		5		64

$C^{(4)}$

a_i \ b_j	2	8	$u_i^{(4)}$	U_i
1	4	0^{1}		-3
9	0^{2}	0^{7}		3
$v_j^{(4)}$				
V_j		5		64

Since $C^{(3)}$ is consistent reduced with the (unique) values of x_{ij} indicated, it follows that $T_c=36$ cost units. $T_0=36+7(1)=8(2)+9(3)=43$ cost units, $T_r=36+7(2)=8(2)+9(3)+7(1)=50$ cost units while $T_h=64$ cost units.

The reverse shipment problem is essentially a special transshipment problem [4, p. 335] in which shipments can be made only from origins to destinations and from destinations to origins.

4. ON ALLOCATIONS WITH SOME NEGATIVE x_{ij}

Once it is recognized that the modification of the Hitchcock problem with the possibility of reshipment, may, in some problems, make possible a total cost which is less than the T_h, the question is raised as to a method for the determination of the minimum total cost. To be sure the method of reduced matrices, without the initial restrictions $x_{ij} \geq 0$ does show (for some problems) that reshipment solutions are useful but, since the method of reduced matrices is designed to produce the best general Hitchcock solutions, even though the formal requirement that $x_{ij} \geq 0$ is not introduced till STEP IV, the process itself, and particularly STEP II which implicitly assumes that $x_{ij} \geq 0$, does not, by itself, guarantee the minimum cost. How can this minimum cost be determined?

First it should be pointed out that there are very, very many allocations with negative $x_{ij} \geq 0$. Essentially the $m \times n$ allocation matrix with marginal totals a_i and b_j fixed has $(m-1)(n-1)$ degrees of freedom so that allocation sets may have as many as $(m-1)\ (n-1)$ independent negative x_{ij} values which are consistent with (2.1) and each may take on any finite negative value. Actually by assigning large absolute values to the negative x_{ij} associated with the larger positive c_{ij}, we can frequently make T_0 less than 0 and in fact, less than any specified negative quantity. However, most of these allocations are of no use in determining reshipment or overshipment solutions since the conditions (3.4) are not commonly satisfied where there are several negative x_{ij} especially if, as is common, all c_{ij} are greater than zero.

If the positions $\alpha\beta$ having $x_{\alpha\beta} < 0$ are known we can use the method of reduced matrices. All that is necessary is to change C to C_- by replacing each $c_{\alpha\beta} > 0$ by $-c_{\alpha\beta}$ and then to find the solution with all x_{ij} positive except for the negative $x_{\alpha\beta}$. This is illustrated in Table 2 with x_{14} negative in a problem used early by Dantzig [3]. The C matrix is first given with the Hitchcock solution values as right superscripts and $T_h = 13$ cost units. This is followed by the C_- matrix, with x_{14} specified to be negative and the reduction by the method of reduced matrices.

This shows that there is a reshipment solution which has a smaller cost, 12 cost units, than the Hitchcock solution. Is there any better one? One approach is to change the sign of each c_{ij}, in turn, to see if any such modified problem, with all other x_{ij} positive, has a smaller T_r. In the illustration above no such smaller T_r results. While this method seems impractical with hand computation, except with trivial problems, it is not unreasonable for problems of moderate size if computers are used. However,

TABLE 2. *Reshipment Solution for Dantzig Illustration with Negative* x_{14}.

C

a_i \ b_j	3	3	3	2	2
1	3	2	1^1	2	3
5	5	4	3^1	-1^2	1^2
7	0^3	2^3	3^1	4	5

$T_h = 13$ cost units.

C_-

a_i \ b_j	3	3	3	2	2	$u_i^{(0)}$	$U_i^{(0)}$
1	3	2	1	−2	3		
5	5	4	3	−1	1		
7	0	2	3	4	5		
$v_j^{(0)}$	0	2	1	−2	1	7	
$V_j^{(0)}$	0	2	1	−2	1		7

$C_-^{(1)}$

a_i \ b_j	3	3	3	2	2	$u_i^{(1)}$	$U_i^{(1)}$
1	3	0	0	0	2	−1	−1
5	5	2	2	1	0		
7	0	0	2	6	4	1	1
$v_j^{(1)}$	−1	−1	1	1		5	
$V_j^{(1)}$	−1	1	2	−1	1		12

$C_-^{(3)}$

a_i \ b_j	3	3	3	2	2	$u_i^{(3)}$	U_i
1	5	2	0^2	0^{-1}	3		−1
5	6	3	1	0^3	0^2		
7	0^3	0^3	0^1	4	3		1
$v_j^{(3)}$							
V_j	−1	1	2	−1	1		12

he possibility exists if several of the c_{ij} are very small, that the best reshipment solution may involve nore than one reverse shipment so a general approach would require the examination of solutions with wo (or more) specified negative x_{ij}—a combinatorial problem of real magnitude. In many cases, how-ver, the optimal reshipment solution is revealed, as it is in the illustration above, by an examination f each case with a possible single negative x_{ij}.

Other methods considered for finding the optimal reshipment solution include a) a modification f the method of reduced matrices designed to produce an initially reduced matrix with smallest R nd with STEP II omitted, and b) the transformations of the completely reduced matrix $C^{(4)}$, or the onsistent reduced matrix $C^{(3)}$, by subtraction from the rows and columns to produce a smaller T_r. As yet, however, these methods are subject to deficiencies similar to those of the method of negative $_{ij}$ above. It appears that a more fundamental approach is in order and, in fact, the most satisfactory olution process depends on a better analysis of the problems in which reshipment solutions are dvantageous.

5. PROBLEMS WITH NONOPTIMAL c_{ij}

In undertaking an efficient determination of useful reshipment solutions, it seems appropriate to iscover, first, those problems which are, or are not, subject to useful reshipment solutions and then o use this information in obtaining optimal reshipment solutions, when appropriate. For this purpose he concept of nonoptimal c_{ij} is introduced. Is the cost of the direct shipment from origin α to destina-ion β greater than the cost from origin α to destination j, plus the cost from destination j to origin i, lus the cost from origin i to destination β? If this is true for some $i, j, i \neq \alpha$ and $j \neq \beta$, as it may well e in Hitchcock problems since shipments from destinations to origins are not permitted, we can say hat the element of the modified problem is "nonoptimal." Alternately, we say $c_{\alpha\beta}$ is nonoptimal if for ome $i \neq \alpha, j \neq \beta$

$$c_{\alpha\beta} > c_{\alpha j} + c_{ji} + c_{i\beta} \qquad \text{or, with } c_{ji} = c_{ij} \tag{5.1}$$

$$c_{\alpha\beta} > c_{\alpha j} + c_{ij} + c_{i\beta}. \tag{5.2}$$

We also have as a measure of nonoptimality of any element $\alpha\beta$ with respect to any other element ij not n row α nor in column β,

$$d_{\alpha\beta,\,ij} = c_{\alpha\beta} - (c_{\alpha j} + c_{ji} + c_{i\beta}) \qquad \text{or, with } c_{ji} = c_{ij} \tag{5.3}$$

$$d_{\alpha\beta,\,ij} = c_{\alpha\beta} - (c_{\alpha j} + c_{ij} + c_{i\beta}). \tag{5.4}$$

f $d_{\alpha\beta,\,\gamma\delta}$ is the largest $d_{\alpha\beta,\,ij}$, we can say, under these conditions, $d_{\alpha\beta,\,\gamma\delta} > 0$ is the measure of non-ptimality of $c_{\alpha\beta}$ with respect to the elements c_{ij}.

If no one of the terms $c_{\alpha j}$, $c_{ij} = c_{ij}$, $c_{i\beta}$ is nonoptimal, then any $c_{\alpha\beta}$ not satisfying (5.2) is said to be optimal." If, under these conditions, (5.2) is satisfied when the greater than sign in replaced by the quality sign, $c_{\alpha\beta}$ is said to be "barely optimal." Of course one or more elements on the right may them-elves be nonoptimal. The definitions above hold generally when all elements on the right, if originally onoptimal, are replaced by elements which are barely optimal. Techniques for accomplishing this

are available in section 7. The measure of nonoptimality is taken to be zero when $d_{\alpha\beta,\,\gamma\delta}$ is negative. For example, Table 1, we know that $c_{11}=1$ cost unit, $c_{12}=2$ cost units, $c_{21}=3$ cost units are all optimal as each is the smallest element in some line. But $c_{22}=8$ cost units $> c_{21}+c_{11}+c_{12}=3+1+2=6$ cost units and the measure of nonoptimality is zero for c_{11}, c_{12}, and c_{21} but $d_{22,\,11}=2$ cost units for c_{22}. And in Table 2 with optimal $c_{24}=-1$, $c_{14}=-2$, $c_{13}=1$ cost units, it is seen that $c_{23}=3>c_{24}+c_{14}+c_{13}=-1+2+1=2$ cost units with $d_{23,\,14}=1$ cost unit. In the first illustration c_{22} is nonoptimal for any value greater than six cost units and in the second illustration c_{23} is nonoptimal for any value greater than two cost units.

In the Hitchcock problem there is no need to consider nonoptimal elements since shipments from destinations to origins are not permitted. As a result many problems, and noticeably those used to illustrate the details of methods proposed by various authors [1], [2], [3], [10] have one or more cost elements which are nonoptimal in the sense used here.

6. RESHIPMENT SOLUTIONS AND NEGATIVE x_{ij}

The definitions above show that for nonoptimal $x_{\alpha\beta}$, there is an indirect route, involving at least one shipment from some destination to some origin, with $T_r < T_h$. Once it is recognized that such a solution may be identified by the negative x_{ij}, we can say that for every nonoptimal $c_{\alpha\beta}$ of the Hitchcock solution there is a smaller total cost featuring some $x_{\gamma\delta}$. However, not every negative $x_{\gamma\delta}$ indicates a nonoptimal $c_{\alpha\beta}$. Thus in Table 1 if c_{22} is changed to the (optimal) four cost units the method of reduced matrices leads to the completely general solution indicated by

$$x_{11}=x_{11}, \qquad x_{12}=1-x_{11}, \qquad x_{21}=2-x_{11}, \qquad x_{22}=7+x_{11}, \tag{6.1}$$

so the Hitchcock solutions are, with $x_{11}^{(h)}=0$ and $x_{11}^{(h)}=1$, respectively,

$$x_{11}^{(h)}=0, \qquad x_{12}^{(h)}=2, \qquad x_{21}^{(h)}=2, \qquad x_{22}^{(h)}=7$$

$$x_{11}^{(h)}=1, \qquad x_{12}^{(h)}=0, \qquad x_{21}^{(h)}=1, \qquad x_{22}^{(h)}=8$$

with, in each case, $T_h=36$ cost units. When $x_{11}=2$, x_{12} is negative and $T_r=40$ cost units and this cost increases with larger integral x_{11}.

7. DETERMINATION OF NONOPTIMAL $c_{\alpha\beta}$

Since negative x_{ij} may be interpreted as reshipments and since reshipments are not necessary when the matrix elements are optimal (though alternative reshipments are available when at least one of the elements used in computing the minimum cost is barely optimal), it follows that the negative solutions are not necessary when all cost elements are optimal. We are thus led to a consideration of the possibility of determining the nonoptimal elements, replacing them by elements which are barely optimal, and then producing a Hitchcock solution for the modified problem which has the net cost of the reshipment solution. It is shown in this and the following sections how this can be done.

The immediate discussion is directed toward the objectives just stated and not to overshipment

solutions, though the methods used here, with slight modifications, are applicable to overshipment solutions as is seen in section 12.

It should be emphasized, at the start, that this approach to the reshipment problem by the determination of and replacement of all nonoptimal $c_{\alpha\beta}$ by elements which are barely optimal makes possible a (Hitchcock) solution for any modified problem having the same original cost matrix irrespective of the values of a_i and b_j

The basic approach is recursive in nature. We test each element in turn for optimality using (5.2) for all possible values $i \neq \alpha$, $j \neq \beta$. We waive the question temporarily, of the possible nonoptimality of $c_{\alpha j}$, $c_{ji} = c_{ij}$, $c_{i\beta}$ and determine the measure of nonoptimality with respect to each c_{ij} by using (5.4). A simple technique is available for computing the $d_{\alpha\beta\mu ij}$ for a given $c_{\alpha\beta}$. We replace c_{ij} by $-c_{ij}$ when $i \neq \alpha$ and $j \neq \beta$, take $v_j^{(0)} = c_{\alpha i}$, $u_i^{(0)} = c_{i\beta} - c_{\alpha\beta}$ and subtract to get

$$d_{\alpha\beta,ij} = \bar{c}_{ij} - u_i^{(0)} - v_j^{(0)}. \tag{7.1}$$

We examine all positive values of $d_{\alpha\beta,ij}$ and take the largest of these $d_{\alpha\beta,\gamma\delta}$ as the measure of nonoptimality of $c_{\alpha\beta}$, assuming that the elements $c_{\gamma\delta}$, $c_{\gamma\beta}$, $c_{\alpha\delta}$ are themselves optimal.

For each $c_{\alpha\beta}$ with $d_{\alpha\beta,\gamma\delta} > 0$, the value of $c_{\alpha\beta}$ is replaced by $c_{\alpha\beta} - d_{\alpha\beta,\gamma\delta}$ to obtain a transformed element which is barely optimal. This is used in the test for optimality of additional elements. The process continues until all nonoptimal elements are replaced by elements which are barely optimal.

The test for nonoptimality of $c_{23} = 3$ cost units in Table 2 is shown in Table 3. Values of a_i and b_j are not used. The value of $c_{\alpha\beta}$ tested is boxed for easy identification.

TABLE 3. *Computation of* $d_{23,ij}$ *for Illustration of Table 2*

$\bar{c}_{ij}$ for c_{23}

a_i \ b_j						$u_i^{(0)}$
	−3	−2	1	−2	−3	−2
	5	4	[3]	−1	1	0
	0	−2	3	−4	−5	0
$v_j^{(0)}$	5	4	3	−1	1	

$d_{23,ij}$

−6	−4	0	1	−2
0	0	[0]	0	0
−5	−6	0	−3	−6

It is seen that the measure of nonoptimality, assuming all other elements optimal, is $d_{23,14} = 1$ cost unit. Hence the barely optimal replacement for c_{23} is $c_{23} - d_{23,14} = 3 - 1 = 2$ cost units and the test for optimality of other elements continues with this replacement.

This recursive approach is not too unsatisfactory for problems of appreciable size in which it is desired to test all elements for optimality if programmed on a digital computer, but modifications may be in order for smaller problems or for problems in which some specific set of a_i and b_j is involved. In such cases it may be easier to start with the smaller elements of the matrix. where the optimality is easily verified, and then use these established results in testing somewhat larger elements until a substantial group of optimal elements or of nonoptimal elements replaced by barely optimal elements,

is available. Thus, in Table 2 it can be established that all elements associated with the Hitchcock solution, except c_{23}, are optimal.

Study of the definition of optimality leads us to the following conditions for the common problems in which there is no negative c_{ij}. These can be modified for limited negative c_{ij} as needed. The conditions are not mutually exclusive.

An element is optimal, $c_{ij} \geq 0$, when a) it is the smallest element in its line, b) it is not greater than the sum of the three smallest elements of the matrix, c) it, less than the smallest element elsewhere in its line, is not greater than the sum of the two smallest elements not in its line.

These conditions for optimality are applied in Table 4 to a problem used as an illustration by Balinski and Gomory [1], [6]. The a_i and b_j of their problem are shown but the results are applicable to problems with general a_i and b_j. Each optimal c_{ij} is marked by an upper prescript indicating determination of optimality according to the conditions a), b), c) applied in that order.

TABLE 4. *Application of Conditions for Optimality*

c_{ij}

a_i \ b_j	2	2	3	4	4
4	b_3	5	b_3	a_1	a_1
5	a_2	b_3	b_3	b_3	7
6	a_1	a_1	a_2	a_1	b_2

Application of the conditions shows that all elements are optimal except possibly c_{12} and c_{25}. Formal reduction shows c_{12} is nonoptimal with $d_{12,34} = 2$ cost units, and c_{25} nonoptimal with $d_{25,14} = 2$ cost units. Condition c) is not needed to establish optimality of any optimal element of this problem though it is applicable to each of them.

Another illustration of the use of the conditions is presented in Table 5. The problem is one which Khun used to illustrate the Hungarian method [10], [6]. THE a_i and b_j are all unity in this problem though the conditions are the same for all a_i and b_j. The conditions a), b), c) are again applied in that order and the optimal elements indicated as before.

TABLE 5. *Second Application of Conditions for Optimality*

a_i \ b_j	1	1	1	1
1	c_8	a_7	c_9	c_9
1	b_5	a_2	7	c_8
1	6	a_1	b_4	c_9
1	a_2	b_3	a_2	a_6

All elements are optimal except possibly c_{23} and c_{31}. Formal analysis shows that these are barely optimal so the matrix is composed of optimal elements.

Thus it is possible to determine nonoptimal elements and to replace them, when discovered, by

elements which are barely optimal. We are now in a position to relate the problem of reshipment to the property of possible nonoptimality of the matrix elements.

8. RESHIPMENT SOLUTIONS FOR COST MATRICES WHICH ARE COMPLETELY OPTIMAL

When every element of the cost matrix is optimal, there is no indirect way in which an individual shipment can be made from origin i to destination j at smaller cost. It follows that no collection with specfied a_i and b_j, characteristic of the transportation problem, can result in a smaller sum when C is compeletely optimal even though reshipments are permitted. For example in the illustration of Table 5, since the elements are known to be optimal, it follows at once that the Hitchcock solution with a sum of 17 cost units [10] cannot be improved with the use of reshipments. It should be noted that this statement does not imply that there is no reshipment solution with a total cost as small as that of the Hitchcock solution. If an element of the direct solution, $x_{\alpha\beta}$ is associated with a barely optimal $c_{\alpha\beta}$, then a reshipment solution is just as good. For example with $a_i = 0, 0, 1, 0$ and $b_j = 1, 0, 0, 0$ in the problem of Table 5, the direct solution gives $T = 6$ cost units and so does the reshipment solution with $x_{41} = 1$, $x_{42} = -1$, $x_{32} = 1$.

9. RESHIPMENT SOLUTIONS USING TRANSFORMED ELEMENTS

The replacement of each nonoptimal element by an element which is barely optimal, i.e., the use of the reshipment cost as a substitute for the direct cost when this is not optimal, makes possible the restatement of the problem in terms of direct shipments with the optimal elements and transformed barely optimal elements as costs. Thus the solution is reduced to a Hitchcock solution with the additional step of transforming each $x_{\alpha\beta}^{(h)} > 0$ associated with nonoptimal $c_{\alpha\beta}$ to the appropriate reshipment solution, i.e., by deleting $x_{\alpha\beta}^{(h)}$ from the $\alpha\beta$ position, adding it in the $\alpha\delta$ and $\gamma\beta$ positions, and adding its negative in the $\gamma\delta$ position. The general solution of the problem of Table 1, with the nonoptimal $c_{22} = 8$ cost units replaced by the barely optimal six cost units, by using the method of reduced matrices, is indicated in Table 6.

This gives the most general Hitchcock solution for the transformed matrix as $x_{12}^{(h)} = 1$, $x_{21}^{(h)} = 2$, $x_{22}^{(h)} = 7$. Replacing the seven units in the 22 position by adding seven units to the 12 and 21 positions and substracting seven units from the 11 position, we arrive at the best reshipment solution possible, with a total cost of 50 cost units, which agrees with the result indicated in Section 3.

TABLE 6. *Solutions of Problems of Table 1 Using Transformed* c_{22}

C (with barely optimal c_{22})

a_i \ b_j	2	8	$u_i^{(0)}$	$U_i^{(0)}$
1	1	2	1	1
9	3	6	3	3
$v_j^{(0)}$		1	36	
$V_j^{(0)}$		1	-	36

$C^{(1)}$

a_i \ b_j	2	8	$u_i^{(1)}$	$U_i^{(1)}$
1	0	0		1
9	0	2	2	5
$v_j^{(1)}$	−2		14	
$V_j^{(1)}$	−2	1		50

TABLE 6. *Solutions of Problems of Table 1 Using Transformed* c_{22} – Continued

$C^{(3)}$

a_i \ b_j	2	8	$u_i^{(3)}$	$U_i^{(3)}$
1	2	0^1		1
9	0^2	0^7		5
$v_j^{(3)}$				
$V_j^{(3)}$	−2	1		50

reshipment

a_i \ b_j	2	8
1	1^{-7}	2^8
9	3^9	8

10. PROBLEMS WITH RESTRICTED RESHIPMENTS

The nature of the solution indicated in the previous section is such that it is immediately available for problems in which restrictions are placed on the size of the reshipment. In many problems, presumably, such restrictions are not pertinent but a slight modification makes the technique available when they are. We simply transform as many of the negative $x_{\gamma\delta}$ as are permitted by the restriction on the size of the reverse shipment at $\gamma\delta$, complete the allocation with the Hitchcock assignments, and compute the new T_r. For example in the reshipment solution of Table 6, if not more than five parcels can be shipped from destination one to origin one, we have $x_{11}=-5$, $x_{12}=6$, $x_{21}=7$, $x_{22}=2$ with $T_r=54$ cost units.

11. OVERSHIPMENT SOLUTIONS RELATED TO RESHIPMENT SOLUTIONS

In every case in which $T_r < T_h$, it follows from (3.3) that $T_0 \leq T_r$ and overshipment, when permitted, is useful in making the cost smaller. It is only necessary, to replace the negative x_{ij} of the reshipment solution by zero and to make the calculation for T_0. Thus an overshipment solution for the problem of Table 6 is $x_{12}=8$ cost units, $x_{21}=9$ cost units with $T_0=8(2)+9(3)=43$ cost units, as shown in section 3. In the problem of Table 2, $x_{14}=-1$ is replaced by zero and $T_0=3(0)+3(2)+1(3)+3(-1)+2(1)+2(1)=10$ cost units.

The question of restricted overshipments may be important. Though a distributor may be willing to receive a few more parcels than ordered, he may not be willing to accept an overshipment of appreciable size. However, the solutions of the sections above are immediately adaptable. One simply uses the reshipment answers with the stated restrictions on overshipments as far as is possible and the Hitchcock answers for the rest. For example the value of T_0, with the number of all overshipments limited to five, in the problem of Table 1, is $6(2)+7(3)+2(8)=49$ cost units.

12. GENERAL OVERSHIPMENT SOLUTIONS, $c_{ij} \geq 0$

As compared with reshipment, which has T_r smaller than T_h only when there are indirect routes with smaller cost than direct routes, i.e., when one or more $d_{\alpha\beta,ij} \geq 0$, there may be overshipment solutions with T_0 smaller than T_h when all c_{ij} are optimal. It is the objective of this section to establish techniques which provide such overshipment solutions. This presentation is more condensed than that of the sections above since the techniques for the solution of the general overshipment problem are very similar to those for the solution of the reshipment problem.

First it should be noted that, in the technique of the previous sections, we could replace $c_{\gamma\delta}$ by zero instead of replacing negative $x_{\gamma\delta}$ by zero. More generally, we can replace $c_{ij}=c_{ji}$, $i\neq\alpha, j\neq\beta$ by zero and get, as a measure of overshipment potentiality of element i, j,

$$e_{\alpha\beta,\,ij}=c_{\alpha\beta}-(c_{\alpha j}+c_{i\beta}). \tag{12.1}$$

If, and only if, this quantity is positive, any shipment of $x_{\alpha\beta}>0$ parcels by route $\alpha\beta$ can be replaced, with smaller net cost by a dual shipment—one through αj and one through $i\beta$. Under these conditions, which result in an overshipment, $T_0<T_h$.

We hence calculate, for each $\alpha\beta$, the measure of overshipment potentiality for each alternate element $i\neq\alpha$ and $j\neq\beta$. We say $i=\gamma$ and $j=\delta$ for any element with the greatest overshipment potentiality. To determine this measure we follow the technique for determining the measure of nonoptimality except that c_{ij} is replaced by zero instead of by $-c_{ij}$. Thus to obtain a measure of overshipment potentiality for c_{22} in Table 1 we have

TABLE 7. *Computation of Overshipment Potentiality*

C_0

0 3	2 8	−6
3	8	

$e_{22,\,ij}$

3	0
0	0

and the measure of overshipment potentiality is $e_{22,\,11}=3$ cost units.

Generally the matrix adjusted for possible overshipments has $c_{\alpha\beta}-e_{\alpha\beta,\,\gamma\delta}$ in place of $c_{\alpha\beta}$, for which we obtain the general Hitchcock solution. This is illustrated in Table 8

TABLE 8. *Overshipment Solution From Modified Cost Matrix*

C(modified)

a_i \ b_j	2	8	$u_i^{(0)}$	$U_i^{(0)}$
1 9	1 3	2 5	3	3
$v_j^{(0)}$		2	43	
$V_j^{(0)}$		2		43

$C^{(4)}$

a_i \ b_j	2	8	$u_i^{(4)}$	$U_i^{(4)}$
1 9	1 0^2	0^1 0^7		3
$v_j^{(4)}$				
$V_j^{(4)}$		2		43

The general Hitchcock solution of this modified problem is $x_{12}^{(h)}=1$, $x_{21}^{(h)}=2$, $x_{22}^{(h)}=7$ with $T_h=43$ cost units. The replacement of the shipment of seven parcels x_{22} be seven parcels x_{12}, and x_{21} leads to the overshipment solution $x_{12}=8$, $x_{21}=9$ with $T_0=8(2)+9(3)=43$ cost units as indicated. In this illustration the overshipment is related to the equivalent c_{22} term, but the technique is also applicable to cost matrices with no equivalent c_{ij}. For example, a problem which is like the above, except that c_{22} is

the (optimal) six cost units, has the overshipment solution given in Table 8, though the Hitchcock solution gives $T_h = 50$ cost units.

The less formal argument here is limited to the usual cases in which $c_{ij} \geq 0$. The treatment is simpler than that of the reverse shipments above.

It is seen at once that there is no useful overshipment resulting from assigning $x_{\alpha\beta} > 0$ to $c_{i\beta}$ and $c_{\alpha j}$ instead of to $c_{\alpha\beta}$ when

$$c_{\alpha\beta} \leq c_{i\beta} + c_{\alpha j} \qquad \text{for all } i, j \tag{12.2}$$

for then the contribution (to the total sum) from $x_{\alpha\beta}$, $(c_{i\beta} + c_{\alpha j})x_{\alpha\beta}$ versus $c_{\alpha\beta}x_{\alpha\beta}$, is not decreased by the overshipment, but it is always decreased when (12.2) is not satisfied and $x_{\alpha\beta} > 0$. Since $c_{i\beta}$ is an element of column β and $c_{\alpha j}$ is an element of row α, we consider only the elements of column β and row α. Also (12.2) is satisfied when $i=a$, when $j=b$, and of course when $i=\alpha$, $j=\beta$ so a necessary condition that (12.2) is not satisfied is then $i \neq \alpha$ and/or $j \neq \beta$.

If i^* is used to indicate an element which is the smallest in column β, and j^* to indicate an element which is the smallest in row α, then (12.2) gives

$$c_{i^*\beta} \leq c_{i^*\beta} + c_{\alpha j^*} \leq c_{i\beta} + c_{\alpha j} \tag{12.3}$$

so a single application is sufficient for each $c_{\alpha\beta}$ using $c_{i^*\beta}$ and $c_{\alpha j^*}$. It is very easy to apply (12.3) when $c_{i^*\beta}$ and $c_{\alpha j^*}$ are marked in some way, since, when (12.3) is not satisfied, $c^*_{\alpha\beta} = c_{i^*\beta} + c_{\alpha j}$ is formed by adding the smallest elements in the lines which intersect at α, β.

The "overshipment cost matrix", C^*, is formed by replacing $c_{\alpha\beta}$ by (the numerical value of) $c^*_{\alpha\beta}$ when (12.3) is not satisfied. The (numerical) values of i^*, j^* are indicated by subscripts of the (numerical) values of $c^*_{\alpha\beta}$ so that the information for replacing the $x_{\alpha\beta} > 0$ to form the overshipment is immediately available. Thus C^* for Table 1 is $\begin{bmatrix} 1 & 2 \\ 3 & 5_{11} \end{bmatrix}$ since the only element not satisfied by (12.3) is $\alpha\beta = 22$ with $i^* = 1$ indicating the first row element of column 2 and $j^* = 1$ indicating the first column element of row 2. Then any $x_{22} > 0$ resulting from a Hitchcock solution of the modified problem is added to x_{12} (since $i^* = 1$) and to x_{21} (since $j^* = 1$) to obtain the overshipment assignment, as shown in the discussion under Table 8.

It is usually very easy to find an overshipment solution when a Hitchcock solution is available and it is known that there is a solution element—an element of C which is associated with a positive assignment—which does not satisfy (12.3). We simply use the numerical values of i^*, j^* as subscripts of the numerical value of the solution element and form the overshipment solution as indicated above. Thus in Table 1 we have

$$\begin{matrix} 1 & 2^1 \\ 3^2 & 8^7_{11} \end{matrix}$$

so the overshipment solution is $x_{12} = 8$, $x_{21} = 9$ with $T_0 = 43$ cost units.

A less trivial overshipment problem is presented in Table 9 in which the cost units for shipments between various cities of the United States are distances rounded to the nearest 100 miles. In such a case all c_{ij} are optimal, as may be verified by the reshipment techniques above, and no reshipment exists for which $T_r < T_h$. However, there is still the possibility of useful overshipment solutions.

The first matrix of Table 9 shows not only the a_i, the b_j, and the c_{ij}, but also the general Hitchcock solution as found by the method of reduced matrices [6, p. 83]. The net row and column subtractions, U_i and V_j, and the value of $T_h = 1{,}643$ cost units are given in the right and bottom margins.

Determination of the solution elements which do not satisfy (12.3) is accomplished in a minute or two. Inspection shows that they are all the smallest elements in their columns with the exception of $c_{24}=5$, $c_{25}=5$, and $c_{36}=19$. It is also immediately apparent that $c_{24} > c_{14}+c_{21}$, $c_{25}=c_{21}+c_{35}$, and $c_{36} > c_{35}+c_{46}$. The proper subscripts are inserted for c_{24} and c_{36} and the indicated overassignments are made to get $x_{14}=92$, $x_{21}=58$, $x_{35}=67$, $x_{46}=79$ which with the unchanged $x_{22}=11$, $x_{25}=19$, and $x_{33}=32$ is an overshipment of $10+36=46$ parcels and with a total cost of only $T_0=1{,}201$ cost units.

This assignment, which decreases the cost by $10(5-2-2)+36(19-3-4)=442$ cost units, more

TABLE 9. *Overshipment Solution for a 4×6 Problem with Optimal Elements*

Origin i \ Destination j		Buffalo	Chicago	New Orleans	New York	St. Louis	San Francisco	
	b_j / a_i	48	11	32	92	50	79	U_i
Boston	82	5	9	16	2^{82}	12	32	−3
Cleveland	88	2^{48}	3^{11}	11	5^{10}_{11}	5^{19}	26	
Kansas City	99	10	6	8^{32}	12	3^{31}	19^{36}_{45}	−2
Los Angeles	43	26	21	19	20	19	4^{43}	−17
	V_j	2	3	10	5	5	21	1,643
	312	48	11	32	92	50	79	U_i^*
Boston	82	4_{24}	5_{24}	10_{34}	2^{82}	5_{34}	6_{14}	0
Cleveland	88	2^{48}	3^{11}	10_{31}	4^{y}	5	6^{29-y}_{41}	2
Kansas City	99	5_{25}	6_{25}	8^{32}	5^{10-y}_{15}	3^{50}	7^{7+y}_{45}	3
Los Angeles	43	6_{26}	7_{26}	12_{36}	6_{16}	7_{36}	4^{43}	0
	V_j^*	0	1	5	2	0	4	1,144
	358	77	11	32	92	67	79	
Boston	92	5	9	16	2^{92}	12	32	
Cleveland	88	2^{77}	3^{11}	11	5	5	26	
Kansas City	99	10	6	8^{32}	12	3^{67}	19	
Los Angeles	79	26	21	19	20	19	4^{79}	

than a quarter of the Hitchcock cost, can be computed from the Hitchcock solution in less than 5 minutes.

The elimination of the assignment at c_{24} results in a decrease of only 10 cost units, but at c_{36} the decrease is 432 cost units. An examination of the reason for this saving may enable us to see more clearly how the overshipment works. It is obviously cheaper to ship 36 parcels from Kansas City to St. Louis and also to ship 36 parcels from Los Angeles to San Francisco than it is to ship, according to the Hitchcock solution, 36 parcels all the way from Kansas City to San Francisco. The result, however, is a surplus of 36 parcels at St. Louis and a depletion of 36 (extra) parcels from the stock at Los Angeles.

Some may feel that, while a few additional parcels may be acceptable, 36 is too many. As in the reverse shipment problem, adjustment is easily made for a bounded overshipment. We simply make as many allocations as possible according to the overshipment bounds and complete the allocation with the Hitchcock solution. Thus if only 14 surplus parcels are acceptable at St. Louis, the shipment is made with $x_{35}=31+14=45$ parcels, $x_{36}=36-14=22$ parcels, $x_{46}=43+14=57$ parcels. Then the T_0 is smaller than T_h by $14(19-3-4)=168$ cost units which is still more than one-tenth of the Hitchcock cost.

It must be kept in mind that these results hold for the a_i and b_j specified. Other specified a_i and b_j may tend to other overshipment solutions.

Although the technique above can usually be applied in a few minutes, once the Hitchcock solution is known, there is no guarantee that it gives the best overshipment solution. For this purpose a more fundamental approach is necessary. The complete C^* matrix is computed with subscripts in all terms not satisfied by (12.3). A general Hitchcock solution is then found for the C^* matrix. (If a particular solution is used, there can be no guarantee that the resulting overshipment solution is general.) This solution is then applied to the C matrix and overshipments introduced as the subscripts indicate.

The second matrix of Table 9 corresponds to that of the first matrix except that the values are appropriate to C^*. The third matrix shows the optimal overshipment solution with an overshipment of $10+29+7=46$ parcels and with $T_r=1{,}144$ cost units. This is 499 cost units less than T_h and represents a saving of approximately 30 percent.

The treatment of the overshipment problem in [9] features the determination of those overshipments having maximum reduction in cost. It was written months later than this section and shows modified techniques and, in some respects, solutions for more general problems.

13. PROBLEMS WITH SOME IMPOSSIBLE DIRECT SHIPMENTS

The method of reduced matrices is also very helpful in solving problems in which, for various reasons, some direct shipments are impossible or impractical. The method of reduced matrices in addition to being general, simple, easy and direct, enables one to discover as the solution process progresses whether or not an actual solution exists so that a prior determination of the existence of a solution, which may be as difficult as the determination of the solution itself, is not required. One simply replaces every impossible c_{ij}, as is conventional, by ∞ or $\times$ and applies the method of reduced matrices, with the modification that no $\times$ value can be used to provide the $u_i^{(t)}$ or the $v_j^{(t)}$. The process leads

either to a solution or to a situation in which the components can be connected only with the subtraction of some $\times$ value, which is impossible. For some of these problems it is possible to obtain a Hitchcock solution while in others such a solution is not possible though a solution may be obtained if negative x_{ij} (reshipments) are permitted. This type of problem is illustrated in Table 10 in which the solution of the problem of Table 1, with c_{22} replaced by $\times$, is worked with the method of reduced matrices. It is seen at once from $C^{(3)}$ that there is a reshipment solution with $T_r = 8(2) + 9(3) + 7(1) = 50$ cost units but no Hitchcock solution, since one cannot subtract $\times$ either in row 2 or in column 2. A second illustration is presented in Table 11.

TABLE 10. *A Simple Illustration with Impossible* c_{22}

C

a_i \ b_j	2	8	$u_i^{(0)}$
1 9	1 3	2 $\times$	1 3
$v^{(0)}{}_j$		1	36

$C^{(3)}$

	2	8	U_i
1	0^{-7}	0^{8}	1
9	0^{9}	$\times$	3
V_j		1	36

TABLE 11. *Reduced Matrix Solution of a Problem with Many Inadmissible Squares* [4, p. 331]

C

a_i \ b_j	14	7	5	17	$u_i^{(0)}$
7	$\times$	3	3	$\times$	
25	2	$\times$	19	6	
8	7	$\times$	4	$\times$	1
3	8	$\times$	$\times$	5	
$v_j^{(0)}$	2	3	3	5	157

$C^{(1)}$

b_j	14	7	5	17	$u_i^{(1)}$
7	$\times$	0	0	$\times$	
25	0	$\times$	16	1	1
8	4	$\times$	0	$\times$	
3	6	$\times$	$\times$	0	
$v_j^{(1)}$	-1				11

$C^{(2)}$

	14	7	5	17	$u_i^{(2)}$
7	$\times$	0	0	$\times$	
25	0	$\times$	15	0	
8	5	$\times$	0	$\times$	5
3	7	$\times$	$\times$	0	
$v_j^{(2)}$			-5		15

$C^{(4)}$

	14	7	5	17	U_i
7	$\times$	0^{7}	5	$\times$	
25	0^{11}	$\times$	20	0^{14}	1
8	0^{3}	$\times$	0^{5}	$\times$	6
3	7	$\times$	$\times$	0^{3}	
V_j	1	3	-2	5	183

14. TRANSSHIPMENT SOLUTIONS WITH REDUCED MATRICES

The possibility of better indirect routes than direct routes is recognized in the statement of the transshipment problem. The general transshipment model, with diagonal supply variables as described by Dantzig [4, p. 339], allows for possible shipments to and from any point. A basic tableau is formed by listing each point as origin and destination with i and j taking on $m+n$ values. Then c_{ij} is the direct cost of shipping from point i to point j with $c_{jj}=0$ and if reverse costs are the same as direct costs, $c_{ji}=c_{ij}$. The model is general enough to cover the case of inadmissible squares in which case shipments may be made by indirect routes. With x_{jj} the amount transshipped through point j, we seek nonnegative x_{ij} which satisfy [4, p. 340]

$$\sum_{i\neq j} x_{ij}-x_{jj}=b_j \qquad j=1, 2, \ldots, m+n$$

(14.1)

$$\sum_{k\neq j} x_{jk}-x_{jj}=a_i \qquad i=1, 2, \ldots, m+n$$

so as to minimize $\Sigma_i\Sigma_j c_{ij}x_{ij}$ where commonly many of the a_i and b_j are zero.

We note first that the method of reduced matrices with slight modification is applicable to the general transshipment problem with the solution properties that nondiagonal terms must be nonnegative and diagonal terms nonpositive. Also if a_i or b_j is zero, the value $a_iu_i^{(t)}$, or $b_jv_j^{(t)}$, is zero no matter what the value of $u_i^{(t)}$, or $v_j^{(t)}$, so a constant may be subtracted from any of these rows or columns without changing the current value of T. Since $c_{jj}=0$, the original matrix is initially reduced and the first STEP consists of introducing at least one additional zero term into every line with $a_i\neq 0$ or $b_j\neq 0$ because the positive allocations must be made to nondiagonal terms. The diagonal term must remain zero to preserve the transshipment property so the quantity subtracted from any row (column) must be the negative of the quantity subtracted from the corresponding column (row). The process continues with STEP II until the matrix is reduced grouped, exclusive of the diagonal terms, and then with STEP III until a consistent solution is reached. If this solution has negative nondiagonal elements, STEP IV should be applied to eliminate them.

The first transshipment problem we consider is that in which, in the language of the transportation problem, shipments can be made from origins to destinations and from destinations to origins, but not between origins nor between destinations. In the notation of the transshipment model all such nonshipments are indicated by inadmissible squares. If there are not too many inadmissible squares, the method of the section above yields a solution. Otherwise the method leads to a situation in which $G_k^{(2)}$ can be joined only with the use of inadmissible squares, and hence there is no solution.

However, any transshipment problem of this type is also subject to a reshipment solution as described in the previous sections. Since the solution process is both simpler and easier with the $m\times n$ cost matrix of the transportation problem rather than with the $(m+n)\times(m+n)$ matrix, with many inadmissible squares of this transshipment problem, and since this method also gives answers which have immediate implications for overshipment, this approach is recommeded.

The solution of the more general transshipment problem with shipments available between all points, except possibly for certain inadmissible squares, is illustrated by the reduced matrix solution of a problem by Dantzig [4, p. 339] in Table 12.

TABLE 12. *Solution of a General Transshipment Problem with Reduced Matrices*

$C^{(1)}$

a_i \ b_j	0	4	0	0	0	4	0	2	$u_i^{(1)}$
7	0	9	9	12	×	×	×	×	9
7	9	0	8	×	9	12	×	×	
0	9	8	0	8	9	13	×	×	
0	12	×	8	0	×	×	10	×	
3	×	9	9	×	0	6	×	13	6
0	×	12	13	×	6	0	5	8	
0	×	×	×	10	×	5	0	9	
0	×	×	×	×	13	8	9	0	−7
$v_j^{(1)}$	−9				−6			7	95

$C^{(3)}$

a_i \ b_j	0	4	0	0	0	4	0	2	$u_i^{(3)}$
7	0	0	0	3	×	×	×	×	12
0	18	0	8	×	15	12	×	×	12
0	18	8	0	8	15	13	×	×	12
0	21	×	8	0	×	×	10	×	9
3	×	3	3	×	0	0	×	0	
0	×	12	13	×	12	0	5	1	
0	×	×	×	10	×	5	0	2	
0	×	×	×	×	26	15	16	0	
$v_j^{(3)}$	−12	−12	−12	−9					36

$C^{(4)}$

a_i \ b_j	0	4	0	0	0	4	0	2	U_i
7	0	0^{7}	0	0	×	×	×	×	21
0	18	0^{-3}	8	×	3	0^{3}	×	×	12
0	18	8	0	5	3	1	×	×	12
0	24	×	11	0	×	×	1	×	9
3	×	15	15	×	0	0^{1}	×	0^{2}	6
0	×	24	25	×	12	0	5	1	
0	×	×	×	19	×	5	0	2	
0	×	×	×	×	26	15	16	0	−7
V_j	−21	−12	−12	−9	−6			7	131

The subtractions from the first matrix are made so as to introduce enough zero terms in every line with $a_i > 0$ or $b_j > 0$ so that the resulting matrix is reduced grouped. The resultant L_k corresponding to the $G_k^{(2)}$, $K = 4$, are

$$L_1 = R_1 + R_2 + R_3 - C_1 - C_2 - C_3 = 3,$$

$$L_2 = R_4 - C_4 = 0,$$

$$L_3 = R_5 + R_6 + R_8 - C_5 - C_6 - C_8 = -3,$$

$$L_4 = R_7 - C_7 = 0.$$

Examining $\langle R_1, R_2, R_3 \rangle \cap \langle C_5, C_6, C_8 \rangle$ we see $c_{12}^{(3)} = 12$ cost units should be subtracted from R_1, R_2, R_3 (with associated subtraction of -12 cost units from (C_1, C_2, C_3) and -9 cost units from C_4). The increase in the reduction is $12(7-4) = 36$ cost units. The matrix resulting from these subtractions is consistent reduced, with $K = 2$

$$L_1 = \sum R_i - R_7 - \left(\sum C_j - C_7\right) = 0$$

$$L_2 = R_7 - C_7 = 0.$$

Actually this matrix is completely reduced (in the transshipment sense with all diagonal x_{ij} nonpositive and all nondiagonal x_{ij} nonnegative) with the solution indicated by the right superscripts in the last matrix of Table 12. The cost is 131 cost units, as shown by Dantzig.

Another variation of the general transshipment problem is the best route problem [4, p. 36]. This is a very special transshipment problem with all $a_i = 0$ except $a_i = 1$ for some point I, $b_j = 0$ except $b_j = 1$ for some point J. Commonly there are many inadmissible squares i.e., only a relatively few direct routes are available. The general method of reduced matrices is useful. One modification, recommended when $m+n$ is large is the subtraction from row I of the value of the largest specified positive element in that row followed by the subtraction from column J of the value of the largest resulting available positive element in that column, with all the associated subtractions necessary to keep the diagonal terms zero and the nondiagonal terms nonnegative. This initial reduction results in a matrix in which the L_k which features R_I also features C_I and all the other rows (and corresponding columns) for which the distances from point I are specified, and a similar L_2 for column J. A problem of Dantzig [4, p. 361] is used as an illustration of the method in Table 13. The problem is the shipment of a parcel from Los Angeles to Boston where the shipment must be made through intermediate cities with the cost of each intercity shipment specified. With the specified costs shown in the $C^{(1)}$ matrix, one subtracts the largest specified c_{Ij} (85 cost units) from the first row (with associated subtractions) and then the largest resulting specified element from the last column (58 cost units) with associated subtractions. The $C^{(2)}$ matrix results with

$$L_1 = R_1 + R_2 + R_3 + R_4 - C_1 - C_2 - C_3 - C_4 = 1,$$

$$L_2 = R_7 + R_8 + R_9 - C_7 - C_8 - C_9 = -1,$$

$$L_3 = R_5 - C_5 = 0,$$

$$L_4 = R_6 - C_6 = 0.$$

TABLE 13. *Best Route Problem With Reduced Matrices*

$C^{(1)}$

		Los Angeles	Portland	Salt Lake City	Dallas	Kansas	Memphis	Chicago	Washington	Boston	$u_i^{(1)}$
		0	0	0	0	0	0	0	0	1	
Los Angeles	1	0	58	43	85	×	×	×	×	×	85
Portland	0	58	0	48	×	×	×	130	×	×	27
Salt Lake City	0	43	48	0	75	66	×	85	×	×	42
Dallas	0	85	×	75	0	29	28	×	×	×	
Kansas City	0	×	×	66	29	0	27	29	62	×	27
Memphis	0	×	×	×	28	27	0	32	53	×	27
Chicago	0	×	130	85	×	29	32	0	40	58	
Washington	0	×	×	×	×	62	53	40	0	25	− 33
Boston	0	×	×	×	×	×	×	58	25	0	− 58
	$u_j^{(1)}$	− 85	− 27	− 42		− 27	− 27		33	58	143

$C^{(2)}$

	b_j / a_i	0	0	0	0	0	0	0	0	1	$u_i^{(2)}$
Los Angeles	1	0	0	0	0	×	×	×	×	×	43
Portland	0	16	0	63	×	×	×	103	×	×	43
Salt Lake City	0	86	33	0	33	24	×	43	×	×	43
Dallas	0	170	×	117	0	29	28	×	×	×	43
Kansas City	0	×	×	108	29	0	27	29	29	×	19
Memphis	0	×	×	×	28	27	0	32	20	×	15
Chicago	0	×	157	127	×	29	32	0	7	0	
Washington	0	×	×	×	×	95	86	73	0	0	
Boston	0	×	×	×	×	×	×	116	50	0	
	$v_j^{(2)}$	− 43	− 43	− 43	− 43	− 19	− 15				43

$C^{(4)}$

	b_j / a_i	0	0	0	0	0	0	0	0	1	U_i
Los Angeles	1	0	0	0^{1}	0	×	×	×	×	×	128
Portland	0	116	0	63	×	×	×	60	×	×	70
Salt Lake City	0	86	33	0^{-1}	33	0	×	0^{1}	×	×	85
Dallas	0	170	×	117	0	5	0	×	×	×	43
Kansas City	0	×	×	137	58	0	23	10	10	×	19
Memphis	0	×	×	×	56	31	0	17	5	×	15
Chicago	0	×	200	170	×	43	47	0^{-1}	7	0^{1}	
Washington	0	×	×	×	×	114	101	73	0	0	− 33
Boston	0	×	×	×	×	×	×	116	50	0	− 58
	V_j	− 128	− 70	− 85	− 43	− 19	− 15		33	58	186

Here the smallest element in $\{R_1, R_2, R_3, R_4\} \cap (C_7, C_8, C_9)$ is 43 cost units so the appropriate subtractions are made. The resulting matrix is then completely reduced in the transshipment sense. The solution indicates the intermediate points as Salt Lake City and Chicago with smallest cost $143 + 43 = 186$ cost units. Thus the method of reduced matrices, with slight modification, is available for the simple and easy determination of general and specific transshipment problems. In addition it should be emphasized that the method of reduced matrices does not require a predetermination of the existence of a solution for the case of inadmissible squares since any impossible problem must lead eventually to a transformation in which the amount to be subtracted is the value of the element in some inadmissible square.

15. CONCLUSION AND SUMMARY

The negative x_{ij} values, which sometimes result from application of the method of reduced matrices (reviewed in some detail in section 2), but not from methods which, in the early stages require $x_{ij} \geq 0$, are of value in finding reshipment and overshipment solutions in cases in which at least one element on the cost matrix is nonoptimal. The technique is then directed toward the determination of such possible nonoptimal elements and their replacement by elements which are barely optimal. With these replacements it is possible to obtain a Hitchcock solution (nonnegative elements) from which useful reshipment solutions, if they exist, are easily obtained. A similar technique is available for overshipment solutions. The paper also shows how various transshipment problems can be solved with reduced matrix methods.

REFERENCES

[1] Balinski, M.S. and R. E. Gomory, "A Primal Method for the Assignment and Transportation Problems," Management Science *10*, 578–593 (1964).

[2] Charnes, A. and M. Kirby, "The Dual Method and the Method of Balas and Ivanescu for the Transportation Model," Cahiers du Centre d'Etudes Recherche Operationnel *6*, 5–18 (1964).

[3] Dantzig, G. B., "Application of the Simplex Method to a Transportation Problem," *Activity Analysis of Production and Allocation*, T. C. Koopmans (ed.) (Wiley, New York, 1951), pp. 359–373.

[4] Dantzig, G. B., *Linear Programming and Extensions* (Princeton University Press, Princeton, N.J., 1963).

[5] Dwyer, P. S. and B. A. Galler, "The Method of Reduced Matrices for a General Transportation Problem," Journal of the Association for Computing Machinery *4*, 308–313 (1957).

[6] Dwyer, P. S., "The Direct Solution of the Transportation Problem with Reduced Matrices," Management Science *13*, 77–96 (1966).

[7] Dwyer, P. S., "Use of Completely Reduced Matrices in Solving Transportation Problems with Fixed Charges," Nav. Res. Log. Quart. *13*, 289–313 (1966).

[8] Galler, B. A. and P. S. Dwyer, "Translating the Method of Reduced Matrices to Machines," Nav. Res. Log. Quart. *4*, 55–71 (1957).

[9] Dwyer, P. S., "A General Treatment of Upper Unbounded and Bounded Hitchcock Problems," Nav. Res. Log. Quart. *21*, 445–464 (Sept. 1974).

[10] Kuhn, H. W., "The Hungarian Method for the Assignment Problem," Nav. Res. Log. Quart. *2*, 83–93 (1955).

[11] Kuhn, H. W. and M. Baumol, "An Approximate Algorithm for the Fixed Charges Transportation Problem," Nav. Res. Log. Quart. *9*, 1–15 (1962).

[12] Mueller-Merbach, H., "An Improved Starting Algorithm for the Ford-Fulkerson Approach to the Transportation Problem," Management Science *13*, 97–104 (1966).

ON SCHEDULING WITH EARLIEST STARTS AND DUE DATES ON A GROUP OF IDENTICAL MACHINES

Charles Tilquin

Université de Montréal

ABSTRACT

This paper examines the (n, m) scheduling problem with n operations distributed among m machines. An algorithm for solving this problem is presented and, gives a good heuristic solution on a wide class of problems. Computational results are reported which demonstrate the efficiency of this approach.

. INTRODUCTION

References [2], [3], [6], and [10] consider the scheduling of n operations on a single machine, with arliest start and due date constraints. A more general problem is solved in [4] where any operation can e performed on any machine of a group of m identical machines. Formally, this problem is to schedule operations on m machines (each operation being performed on any single machine). Operation i has duration d_i (on any machine), it may start at and after any time a_i, and it has to be completed by time . An optimal schedule is any schedule which minimizes the maximum lateness of any operation, that is, ie time by which the completion of operation i exceeds b_i.

The algorithm presented in [4] for this problem is based on an implicit enumeration procedure and ses the labeling type algorithm designed for the preemptive case. Computational results achieved ith this method are satisfactory for "reasonable" size problems, but computer time requirements may ecome quickly prohibitive for larger problems.

The algorithm presented here gives good heuristic solutions with very little computational effort r the same problem considered in [4]. It is especially designed toward the solution of large-size oblems. Its originality lies in the fact that it partitions the operations set among the machines early the solution procedure. This feature cuts drastically the number of solutions considered in a sub-quent enumerative phase. For example, for the simple (four operations—two machines) problem, ere exists 49 solutions. At most, six of these solutions will be examined by the proposed algorithm. ie solution procedure is thus considerably faster and this is an important time gain when the algo-hm is used to compute lower bounds for the implicit enumeration algorithm in complex scheduling oblems [7–9]. The auxiliary problems arising in this case consist in scheduling n operations on m entical machines where each operation has a duration d_i, with earliest start a_i and successor opera-ns requiring an additional processing time q_i on other groups of machines. The problem is to dis-bute the operations among the machines and to determine an order of processing the operations located to each machine minimizing—over all possible distributions and orders—the maximum mpletion time—including the q_i's—over all operations. This auxiliary problem is equivalent (proved [3]) to the earliest start and due date (n operations, m machines) problem defined above. In the follow-g, the "auxiliary" form of the problèm is considered and the problem is referred to as the "(n, m)

auxiliary problem," the expression "relaxed problem" being reserved for the $(n, 1)$ auxiliary problem as in [3].

Our approach for solving the (n, m) auxiliary problem involves four steps. It parallels the algorithm developed in [3] for the relaxed problem:

1. computation of lower bounds
2. calculation of an initial solution
3. test of optimality
4. construction of new solutions.

2. BASIC NOTATIONS AND DEFINITIONS

Let $N=\{1,2,\ldots,n\}$ be the set of operations and $M=\{1,2,\ldots,m\}$ be the set of machines.

Let S_j represent a *partial schedule:* $(k_1, k_2, \ldots, k_i, \ldots, k_{r_j})$ with $r_j \leq n$ and $k_i \in N$ for each i) of r_j operations on machine j. Let t_i denote the time by which operation i is scheduled. Then, the *critical job* of a partial schedule S_j is the job i for which *completion time* $c_i = t_i + d_i + q_i$ is maximum among jobs in S_j. The completion time of the critical job of S_j will be denoted by c_j^* and called the value of S_j.

A *schedule* is a partition of N into m partial schedules $S_1, \ldots, S_m$. The critical job of a schedule S is the job for which completion time is maximum among critical jobs for partial schedules partitioning S. The completion time $U(m) = \max_j \{c_j^*\}$ of the critical job of S is called the value of S. Let S^* denote an optimal schedule, that is one for which $U(m) = U^*(m)$ is minimum over all possible schedules S.

A *partial block* B_{ij} — ith block for partial schedule S_j — is a group of operations scheduled on machine j such that the first operation is scheduled at time $t_i = a_i$ and all the successor operations on machine j are processed without delay but at time $t_i > a_i$.

A *block* B_k — kth block for schedule S — is a group of operations such that the first operation on each machine is scheduled at time $t_i = a_i$ and the relation $A_k \geq D_k$ is true. A_k is the minimum a_i over the set of operations so far unscheduled, and D_k is the maximum of the $t_i + d_i$'s for the set containing the last job scheduled on each machine.

Finally, $\{q_i^*,\ i=1,\ \ldots,\ n\}$ and $\{a_i^*,\ i=1,\ \ldots,\ n\}$ are permutations of the q_i's and a_i's in increasing order, and let $v_1 = \sum_{i=1}^{m} q_i^*$ and $v_2 = \sum_{i=1}^{m} a_i^*$.

3. THE HEURISTIC ALGORITHM

STEP 1: *Computation of lower bounds.*

Two lower bounds for the solution $U^*(m)$ of the (n, m) auxiliary problem are calculated. The first $(LB1)$ is independent of the number of machines and is similar to $lb1$ in [3]. Its calculation is straightforward:

$$LB1 = \max_{i \in N} \{a_i + d_i + q_i\}.$$

The second $LB2(m)$ depends on m, the number of machines. In [3], $lb2$ was computed on the last block of a solution structure arising from the scheduling of the operations by increasing a_i's. It is known that this block structure is that of the optimal solution, and that by scheduling by increasing a_i's, even if not optimal, will always produce this structure if it exists in the optimal solution.

In the multimachines problem, the structure of a solution depends both on a partition of the set of operations and on their order on each machine. For this reason, it has been impossible to define *a priori* a scheduling procedure which reveals necessarily a block structure if it exists in the optimal solution. In the proposed algorithm, $LB2(m)$ is initially calculated over all operations. Later in the execution of the algorithm, if a solution shows a block structure, $LB2(m)$ will be recalculated on the last block of this structure. We define $LB2(m)$ as follows:

$$LB2(m) = \left[\frac{\sum_{i=1}^{n} d_i + v_1 + v_2}{m}\right]$$

$[x]$ denotes the smallest integer greater than or equal to x. This expression of $LB2(m)$ corresponds to the following assumptions:

i. Each machine begins to work as soon as one operation is available for processing and the last operation performed on each machine has one of the m shortest "tails" q_i.

ii. there is no idle period on any machine once it is on.

iii. it is possible to divide the work among the machines in such a way that the last operation on each machine is completed at the same hour, up to a time unit, if the division is not integer.

STEP 2: *Calculation of an initial solution.*

In [10], Schrage has presented a simple procedure for constructing a "good" initial solution for the relaxed problem. The procedure proposed here is inspired of Schrage's algorithm. An initial solution is constructed in the following manner:

Phase i. Initialization

Set $T \equiv (T_1, T_2, \ldots, T_m) = (-1, -1, \ldots, -1)$;

Set $c^* \equiv (c_1^*, c_2^*, \ldots, c_m^*) = (0, 0, \ldots, 0)$;

Set $t = -1$;

The set of unscheduled operations $V = \{1, 2, \ldots, n\}$.

$S \equiv (S_1, S_2, \ldots, S_m) = \emptyset$.

Phase ii. Scheduling of an operation

If there is no job $i \in V$, such that $a_i \leq t$, set $t = \min_{i \in V} a_i$.

Set $J = \{i | i \in V \text{ and } a_i \leq t\}$ and $J^* = \{j | j \in V \text{ and } d_j = \max_{i \in J} d_i\}$.

If $|J^*| = 1$, then the operation $j \in J^*$ is to be scheduled.

If not, choose $j \in J^*$ such that $q_j \geq q_k$ for each $k \in J^*$.

Schedule operation j on machine k, where $T_k = \min_{i \in M} T_i$.

Add operation j to the partial schedule S_k.

Phase iii. Updating the constants.

Set $T_k = \max(T_k, a_j) + d_j$;

Set $c_k^* = \max(c_k^*, T_k + q_j)$;

Set $t = \min_{i \in M} T_i$;

Substract operation j from the set V of uncheduled operations.

If $V = \emptyset$, go to phase (iv). If not, go the phase (ii).

Phase iv. Constructing the initial solution

The schedule $S = (S_1, S_2, \ldots, S_m)$ has been constructed. Consider each machine j independently and its set S_j of operations. Then apply Schrage's algorithm [10] to the m relaxed problems separately to construct new partial schedules S'_j which are permutations of the partial schedules $S_j, j = 1, 2, \ldots, m$. The initial schedule is $S' = (S'_1, S'_2, \ldots, S'_m)$,

STEP 3. *Testing the optimality of the initial solution.*

Let $U'(m)$ be the value of the initial schedule S'. If $U'(m) = LB(m)$, where $LB(m) = \max(LB1, LB2(m))$, then S' is the optimal schedule and the algorithm stops here. When $U'(m) \neq LB(m)$, if S' exhibits a block structure, $LB2(m)$ is recalculated on the last block. If $LB2(m) = U'(m)$, then S' is optimal. In other cases, go to the following step of the algorithm.

STEP 4. *Construction of new solutions.*

If the initial schedule S' is not optimal, the last step in the proposed algorithm is to find an optimal (or near optimal) schedule by calculating, separately, optimal or near optimal solutions for some or all of the m relaxed problems associated with the initial schedule S'. The procedure is as follows:

Phase i. Initialization

Set $(u_1, u_2, \ldots, u_m) = (0, 0, \ldots 0)$.

$R = (r_1, r_2, \ldots, r_m)$ is a permutation of the set of M machines by order of decreasing completion time c_j^* of the critical job of S'_j which is the initial partial schedule on machine j.

Set $j = 1$

Phase ii. Solving a relaxed problem

Select machine r_j and apply the procedure developed in [3] for this one-machine problem of $|S_{r_j}|$ operations. Let u_j be the value of the best solution obtained for this particular relaxed problem.

Phase iii. Testing

If $j = m$, stop the algorithm. A solution for the (n, m) auxiliary problem has been found. The schedule generated is optimal if $LB(m) = \max_j u_j$. If $j < m$, stop the algorithm if there is a machine k, such that $u_k \geq c^*_{r_{j+1}}$. In this case, it is impossible to find a better solution than the present one which is optimal if $LB(m) = \max_{j \in M} u_j$. When $u_k < c^*_{r_{j+1}}$ for all $k \in M$, if $c^*_{r_{j+1}} = LB(m)$, stop the algorithm for the solution is optimal; otherwise, set $j = j + 1$ and go to phase (ii).

4. DISCUSSION

It should be noted that the algorithm has been designed in such a way that the number of the relaxed problems to be solved is as small as possible. Calculations are thus drastically reduced in two ways:

i. by determining immediately the distribution of the operations among the machines.

ii. by solving only a small number of corresponding relaxed problems.

Point (i) implies that we accept eventual loss of the optimal solution at the point of departure of our algorithm. This loss is balanced by the fact that the complex (n, m) auxiliary problem is replaced by m one-machine problems, only some of which need to be solved. Even if the algorithm fails to find an optimum, it will always give lower and upper bounds for the optimum at a minimum calculation cost.

The calculation of an initial solution is a very important phase of the algorithm since it determines the distribution of the operations among the machines. Another way of generating such a solution has been tried. It is a direct generalization of Schrage's procedure [10] for the relaxed problem. It has been

isregarded since it has proved to be better than the proposed algorithm in only 10 percent of the 234 roblems on which the algorithm was tested. The approach chosen results in a good distribution of ıe loads among the machines since it schedules very soon those operations with large d_i which per- ırbate the most the distribution of the loads between the machines. By scheduling these operations arly, the perturbations they cause can be damped by their successor operations.

. EXAMPLE

The following 12-operation example will illustrate the algorithm

i	1	2	3	4	5	6	7	8	9	10	11	12
a_i	0	0	1	5	10	12	15	18	22	26	30	36
d_i	1	10	6	2	19	14	30	9	18	6	10	7
q_i	6	15	4	12	20	4	11	14	15	28	13	10

i. *Lower bounds*

$$LB1 = \min_{i \in N} \{a_i + d_i + q_i\} = 60 \text{ for } i = 10.$$

$$q_i^* = \{4, 4, 6, 10, \ldots\}$$

$$a_i^* = \{0, 0, 1, 5, \ldots\}$$

Number of machines: m	1	2	3	4
$LB2(m)$	136	70	49	41
$LB(m) = \max(LB1, LB2(m))$	136	70	60	60

ii. *Initial solution.*—Consider the two-machine problem. After phase (iii) of the calculation of an ıitial solution, the following schedules are produced:

Machine 1's schedule

i	t_i	b_i	c_i
1	0	1	7
3	1	7	11
4	7	9	21
5	10	29	49
9	29	47	62
11	47	57	70
12	57	64	74
10	64	$\underline{70}$	$\boxed{98}$

Machine 2's schedule

i	t_i	b_i	c_i
2	0	10	25
6	12	26	30
7	26	56	67
8	56	$\underline{65}$	$\boxed{79}$

Consider now the two one-machine relaxed problems (phase iv of the calculation of an initial solution). To each, apply Schrage's algorithm to generate the initial schedule S'.

Relaxed problem 1

i	t_i	b_i	c_i
1	0	1	7
3	1	7	11
4	7	9	21
5	10	29	49
10	29	35	63
9	35	53	68
11	53	63	76
12	63	$\underline{70}$	$\boxed{80}$

Relaxed problem 2

i	t_i	b_i	c_i
2	0	10	25
6	12	26	30
8	26	35	49
7	35	$\underline{65}$	$\boxed{76}$

iii. *Testing the optimality.*—For the two-machine problem, $LB(m) = 70$. So the test for optimality fails since $U'(m) = 80$. But the initial solution is structured in two blocks: operations 1 to 4 belong to the first, the others to the second. A new lower bound $LB2(m)$ is calculated on the last block exactly in the same way as above.

$$q_i^* = \{4, 10, \ldots\}$$

$$a_i^* = \{10, 12, \ldots\}$$

$$LB2(m) = \left[\frac{113 + 14 + 22}{2}\right] = 75$$

The test for optimality fails again.

iv. *Construction of new solutions.*—The first relaxed problem is an eight-operation problem with upper bound 80.

The lower bound is $lb = \min_i q_i + \Sigma_i d_i + \min_i q_i$ [3], where $i = 5, 10, 9, 11, 12$, which are the operations on the last block. Then $lb = 10 + 60 + 10 = 80$ and the present solution of the first relaxed problem is optimal. With this solution, $u_1 > c_2^* = 76$, it is not necessary to consider the second relaxed problem. In conclusion, the better lower bound produced by the method for this (12, 2) auxiliary problem is 75 and the better upper bound is 80.

6. TIMING TEST OF THE ALGORITHM

The algorithm has been tested on a series of 234 randomly generated problems. The test has been designed as in [4] for various values of the parameters n, m, $a_{\max}$, $q_{\max}$, and $d_{\max}$. The a_i's were selected from a uniform distribution between 0 and $a_{\max}$, the d_i's from a uniform distribution between 1 and $d_{\max}$ and the q_i's from a uniform distribution between 0 and $q_{\max}$.

Two sets of problems were tested. The first one is the set tested in [4], the other is a set of 54 500-

operation problems. Results are given in Tables I to III with the following notations:

n, m : numbers of operations and machines.

p : number of problems of one type.

op : number of optimal solutions generated in a set of p problems.

t : value of the median of solution times for a set of p problems. All times are expressed in seconds of CPU time, by using the optimizing FORTRAN compiler on a CDC CYBER 74, and do not include compilation time.

f : degree of easiness of a problem. For each choice of parameters, $f = m \cdot (a_{\max} + 1/2 d_{\max})/n \cdot 1/2 d_{\max}$ measures the difficulty of a problem by relating the average total available time to complete all the operations without delay (i.e., starting at a_i) to the average total time needed to complete all the operations. A value of $f < 1$ indicates that most of the jobs will be available for scheduling at the same time, which means that the number of possible processing orders is large and that the problem is difficult. In the same way, $f = 1$ indicates that a small number of jobs (on an average, only one), will be available for scheduling at one time and $f > 1$ implies the existence of idle periods on the machine and thus easier problems with block structured solutions. Nevertheless, f is far from being a "perfect" indicator of the difficulty of a problem because it does not take into account all aspects of the scheduling problem. For example, the distribution of the operations among the machines does not appear in f, even though it is an important factor in the solvability of a problem.

e : degree of nonoptimality for the set of p problems.

$$e = \frac{\sum_{k=1}^{p} (u_k - l_k)}{1/2 \sum_{k=1}^{p} (u_k + l_k)},$$

where $u_k (l_k)$ is the best upper (lower) bound for the solution of problem k. For an optimal solution $u_k = l_k$, so for a set of p problems, when the p solutions are optimal, $e = 0$. (e is the average maximum deviation from the optimum.) For example, $e = 2$ indicates roughly that the solution value is 102 when the optimum is 100.

Table I contains our main results. It corresponds to 500-operation problems performed, respectively, on 10, 20, 35, and 50 machines. In 34 out of 54 problems, the algorithm found an optimum solution. These results are quite encouraging for large problems. When an optimum solution was not found, the degree of nonoptimality e reached 2 percent in only one case. The median solution times are very low. When the optimum is reached, it is obtained in about 1.5 seconds. In other cases, the search ends after 5 or 6 seconds. The search for a solution was stopped for economic reasons when the solution of a relaxed problem was not found after 10,000 cycles of the algorithm from [3].

Tables II and III present the results corresponding to two sets of problems generated exactly in the same way as sets 1 and 2 in [4]. It can be seen that for set 1, we fail to reach the optimum in twice as many cases as in [4]. This is difficult to explain when, in set 2, results are similar to those of [4], but take 10 to 100 times less computer time.

TABLE I. $n=500$; $a_{max}=150$; $q_{max}=25$

m	d_{max}	p	op	f	e	t
10	10	3	0	0.62	0.66	3.58
	20	3	3	0.32	0.0	3.99
	30	3	2	0.22	0.04	4.39
20	10	5	5	1.24	0.0	1.33
	20	5	0	0.64	1.02	3.44
	30	5	0	0.44	0.63	3.71
35	10	5	5	2.17	0.0	1.40
	20	5	5	1.12	0.0	1.49
	30	5	0	0.77	1.98	6.22
50	10	5	5	3.10	0.0	1.52
	20	5	5	1.60	0.0	1.51
	30	5	2	1.10	0.31	1.69

TABLE II. $q_{max}=15$

	d_{max}	p	op	f	e	t
$m=2$	10	10	7	1.47	0.97	0.01
	15	10	1	1.02	5.14	0.03
$n=15$	20	10	0	0.80	5.24	0.02
$a_{max}=50$	25	10	1	0.67	3.25	0.02
$m=3$	10	10	10	2.52	0.00	0.02
	15	10	9	1.72	0.17	0.02
$n=25$	20	10	6	1.32	1.34	0.02
	25	10	3	1.08	1.63	0.03
$a_{max}=100$	30	10	0	0.92	3.05	0.04

TABLE III. $n=25$; $a_{max}=100$; $d_{max}=25$; $q_{max}=15$

m	p	op	f	e	t
2	10	4	0.72	0.85	0.03
3	10	3	1.08	3.69	0.03
4	10	6	1.44	2.74	0.02
5	10	8	1.80	0.50	0.02
6	10	10	2.16	0.00	0.02
7	10	9	2.52	0.04	0.02
8	10	10	2.88	0.00	0.03
9	10	10	3.24	0.00	0.03
10	10	10	3.60	0.00	0.03

7. CONCLUSION

The algorithm gives lower and upper bounds for the solution of any problem when it cannot reach the optimum solution. The difference between the upper and the lower bounds gives a good idea

of the precision of the method. In most of the cases, e ranges from 0 to 2 percent. It is only in small problems ($n \leq 25$) that the error e happens to be significant: $e = 5$ percent.

However, one can wonder if a nonoptimum solution, even if it is very close to the optimum, is of any value. The answer is yes. We justify that. The algorithm has been designed first to be used in the computation of bounds for multimachine, complex job-shop problems. An important number of auxiliary problems are generated along the solution process of such problems [7, 8]. Auxiliary problems' solutions are used as bounds to cut branches in an implicit enumeration algorithm. The better the solution of an auxiliary problem, the better will be the bound in the enumeration tree. The only justification of bounding and cutting branches is the gain of time. If bounding is too time-consuming—that is if it takes too much time to solve an auxiliary problem—then we lose on one hand what we are gaining on the other. So there must be a compromise and we think our algorithm corresponds to such a compromise: it is very fast and produces satisfactory solutions.

BIBLIOGRAPHY

[1] Balas, E., "Machine Sequencing via Disjunctive Graphs: An Implicit Enumeration Algorithm," Operations Res. *17*, 941–957 (Nov.–Dec. 1969).

[2] Bratley, P., M. Florian, and P. Robillard, "Scheduling with Earliest Start and Due Date Constraints," Nav. Res. Log. Quart. *18*, 511–519 (Dec. 1971).

[3] Bratley, P., M. Florian, and P. Robillard, "On Sequencing with Earliest Starts and Due Dates with Application to Computing Bounds for the (n/m/G/Fmax) Problem," Nav. Res. Log. Quart. *20*, 57–67 (Mar. 1973).

[4] Bratley, P., M. Florian, and P. Robillard, "Scheduling with Earliest Start and Due Date Constraints on Multiple Machines," Nav. Res. Log. Quart. *22*, 165–173 (Mar. 1975).

[5] Charlton, J. M. and C. C. Death, "A Method of Solution for General Machine Scheduling Problems," Operations Res. *18*, 689–707 (July–Aug. 1970).

[6] Dessouky, M. I. and C. R. Morgenthaler, "The One-Machine Sequencing Problem with Early Start and Due Dates," Technical Report, University of Illinois at Urbana-Champaign (Apr. 1972).

[7] Florian, M., C. Tilquin, and G. Tilquin, "An Implicit Enumeration Algorithm for Complex Job-Shop Problems," Département d'Informatique de l'Université de Montréal, Publication No. 102 (Juil. 1972). To appear in the *International Journal of Production Research*, 1975.

[8] Florian, M., P. Trépant, P. and A. McMahon, "An Implicit Enumeration Algorithm for the Machine Sequencing Problem," Management Science: Applications *17*, 782–792 (Aug. 1971).

[9] Roy, B. and B. Sussman, "Les problèmes d'ordonnancement avec contraintes disjonctives", SEMA, Direction Scientifique, document technique, No. 23 (Dec. 1971).

[10] Schrage, L., "Obtaining Optimal Solutions to Resource Constrained Network Scheduling Problems," Center for Mathematical Studies in Business and Economics Reports, University of Chicago (Sept. 1970).

[11] Schrage, L., "Solving Resource Constrained Network Problems by Implicit Enumeration-Non-Preemptive Case," Operations Research *18*, 263–278 (Mar.–Apr. 1970).

AN IMPROVED BRANCH AND BOUND PROCEDURE FOR $n \times m$ FLOW SHOP PROBLEMS*

S. S. Panwalkar

Department of Industrial Engineering
Texas Tech University
Lubbock, Texas

and

A. W. Khan

American Science and Engineering Co.
Houston, Texas

ABSTRACT

This paper considers the classical $n \times m$ flow shop sequencing problem. An improved branch and bound procedure is proposed. Computational experience shows that the proposed procedure is more efficient compared to the existing optimizing procedures.

INTRODUCTION

Consider the classical $n \times m$ flow shop sequencing problem with n jobs and m machines. All jobs ow the same sequence of operations and are available at time zero. No passing is allowed and jobs st be scheduled as early as possible. The objective is to obtain a sequence that will minimize the kespan.

For $m=2$, Johnson's well known algorithm [6] can be used to find an optimal sequence very ciently. For $m=3$ (also for $m=2$), "no passing" assumption is not necessary and several researchers 7, 11] have used the branch and bound technique to obtain an optimal sequence.

For the general case with $m>3$, Brown and Lomnicki [3] proposed an extension of the branch and ınd method [7] with no passing assumption. Several other optimizing methods have also been posed [2, 4, 8, 9, 11–13]. Szwarc [12] refers to these methods as "elimination methods."

In the following, an extension to the branch and bound method [3] will be described. This extension ased on adding a characteristic common to the elimination methods (called elimination rule) to the nch and bound procedure. The proposed procedure was tested for the computational efficiency. ults indicate that this procedure represents an improvement over most existing optimizing pro- ures for the $n \times m$ flowshop problem.

NOTATION

Let $N=\{i \mid i=1, 2, \ldots, n\}$ denote a set of n jobs and $M=\{j \mid j=1, 2, \ldots, m\}$ denote a set of ıachines. Each job is processed on machine 1, then on machine 2, and so on. Let σ, π etc. denote ial schedules where a partial schedule represents a subset of N arranged according to the desired

*Partial support for the research was provided by the National Science Foundation Grant GK 2869.

sequence (σ, π may be empty). If an additional job k is added to the schedule σ, the new schedule wil be denoted by σk. Let P_{ij} be the processing time of job i on machine j and let $t(\sigma, j)$ denote the comple tion time of σ on machine j. Then, from [12]

$$t(\sigma i, j) = \max [t(\sigma i, j-1), \quad t(\sigma, j)] + P_{ij},$$

with $t(\phi, j) = t(\sigma, 0) = 0$.

3. THE PROPOSED BRANCH AND BOUND PROCEDURE

The basic procedure consists of three major steps: calculation of lower bounds, branching, and node elimination. These steps and the logic involved is quite similar to [3] except in the node elimination. Briefly, the procedure involved is as follows.

Calculation of Lower Bounds.—If a node represents a partial schedule σ, $t(\sigma, j)$ for $j = 1, 2, \ldots, m$ is calculated. This in turn is used to calculate the machine based bounds [3, 10].

Branching.—Branch from the node with the smallest value of the lower bound [3].

Node Elimination.—In this step two methods are used.

(A) Set the upper bound equal to the best makespan value among the complete schedules. (Initially the upper bound is set to infinity). Eliminate all nodes having lower bound values greater than or equal to the current upper bound [3].

(B) Let σ denote the partial sequence corresponding to the node selected for branching. Let i, k be any two jobs not included in σ. Then, elminiate node σk if

$$t(\sigma ik, j-1) - t(\sigma k, j-1) \leq t(\sigma ik, j) - t(\sigma k, j) \leq P_{ij} \qquad j = 2, 3, \ldots, m.$$

This represents the elimination rule proposed by Szwarc [12].

4. COMPUTATIONAL RESULTS AND DISCUSSION

The proposed method, the branch and bound method of Brown and Lomnicki [3], and the Szwarc method [12] were programmed in FORTRAN IV and run on IBM 370/145. Several problems were solved and the computation times in seconds (excluding compilation and $I/0$ times) are shown in Table 1. The problems were generated by using random numbers from 0–99 uniform distribution.

Discussion.—The proposed method combines Szwarc's elimination rule with Brown and Lomnicki's bound procedure. Instead, Smith-Dudek [11] or McMahon's [8] elimination rule could have been combined with the branch and bound procedure. Since these rules are quite similar in nature, no significant difference in the performance was observed when these other rules were substituted for solving the test problems.

A method such as Szwarc's [12] does not appear to be as efficient as the branch and bound method [see Table 1]. This is because it does not use any upper and lower bounds for elimination. Several other elimination methods [4, 8, 9, 11, 12] were also tested. The performance of these methods is quite similar to the Szwarc method as these methods also do not use upper and lower bounds. Baker [1] has also done a comparison of existing flowshop algorithms.

TABLE 1. *Average Computation Times (sec) for the Branch and Bound and the Proposed Method*

Size	Number of problems	Szwarc	Branch and bound	Proposed method
5×5	20	0.224	0.185	0.151
5×10	20	0.580	0.400	0.391
6×5	20	1.229	0.690	0.518
6×10	20	5.299	1.370	1.538
7×5	20	6.824	2.356	1.999
7×10	20	52.364	4.996	5.065
8×5	20	49.819	7.458	5.669
8×10	20	(*)	16.320	15.097
9×5	20	(*)	29.391	18.337
9×10	20	(*)	60.885	46.744
10×5	20	(*)	77.937	43.338
10×10	20	(*)	345.844	233.045
12×5	20	(*)	281.585	164.091
12×10	20	(*)	(*)	(*)

*Problems of this size either required excessive computer time or exceed the storage space.

Finally, the computational efficiency of the proposed method which combines the existing methods appears to be substantially higher. Further improvements in the performance will depend upon improved branching procedures, improving the methods of finding upper and lower bounds or improving the computer program.

BIBLIOGRAPHY

[1] Baker, K. R., "A Comparitive Study of Flow Shop Algorithms," Operations Res. *23*, 62–73 (1975).

[2] Baker, K. R., "An Elimination Method for the Flow Shop Problem," Operations Res. *23*, 159–162 (1975).

[3] Brown, A. P. G. and Z. A. Lomnicki, "Some Applications of the 'Branch-and-Bound' Algorithm to the Machine Scheduling Problem," Opl. Res. Q. *17*, 173–186 (1966).

[4] Gupta, J. N. D., "An Improved Combinatorial Algorithm for the Flowshop-Scheduling Problem," Operations Res. *19*, 1753–1758 (1971).

[5] Ignall, E. and Schrage, L., "Application of the Branch and Bound Technique to Some Flowshop Problems," Operations Res. *13*, 400–412 (1965).

[6] Johnson, S. M., "Optimal Two-and-Three Stage Production Schedules with Setup Times Included," Nav. Res. Log. Quart. *1*, 61–68 (1954).

[7] Lomnicki, A. A., "A Branch-and-Bound Algorithm for the Exact Solution of a Three-Machine Scheduling Problem," Opl. Res. Q. *16*, 89–100 (1965).

[8] McMahon, G. B., "A Study of Algorithms for Industrial Scheduling Problems," Ph. D. Thesis, University of New South Wales, Sydney, Australia (1971).

[9] McMahon, G. B., "Optimal Production Schedules for Flow Shops," Jr. of Canadian OR Society *7*, 141–151 (1969).

[10] McMahon, G. B. and P. G. Burton, "Flowshop Scheduling with the Branch and Bound Method," Operations Res. *15*, 473–481 (1967).

[11] Smith, R. D. and R. A. Dudek, "A General Algorithm for the Solution of the *n*-Job, *m*-Machine Sequencing Problem of the Flowshop," Operations Res. *15*, 71–82 (1967).

[12] Szwarc, W., "Elimination Methods in the $m \times n$ Sequencing Problem," Nav. Res. Log. Quart. *18*, 295–305 (1971).

[13] Szwarc, W., "Optimal Elimination Methods in the $m \times n$ Flowshop," Operations Res. *21*, 1250–1259 (1973).

IMPLICIT ENUMERATION BASED ALGORITHMS FOR POSTOPTIMIZING ZERO-ONE PROGRAMS*

Christopher J. Piper

University of Western Ontario
London, Canada

and

Andris A. Zoltners

University of Massachusetts
Amherst, Massachusetts

ABSTRACT

It has been shown by G. Roodman that useful postoptimization capabilities for the 0–1 integer programming problem can be obtained from an implicit enumeration algorithm modified to classify and collect all fathomed partial solutions. This paper extends the the approach as follows: 1) Improved parameter ranging formulas are obtained by higher resolution classification criteria. 2) Parameters may be changed so as to tighten the original problem, in addition to relaxing it. 3) An efficient storage structure is presented to cope with difficult data collection task implicit in this approach. 4) Finally, computer implementation is facilitated by the elaboration of a unified set of algorithms.

INTRODUCTION

Roodman [9] has described a new procedure for postoptimizing zero-one integer programs. Based upon implicit enumeration, it employs information collected during the optimization of the initial problem. The information consists of a list of examined partial solutions, and the way they were fathomed by the algorithm.

Roodman observed that, whenever search along a branch of the implicit enumeration tree is terminated, the fathomed partial solution, s, can be *attributed* to a constraint, k. Unless constraint k is somehow relaxed, s and all its completions will remain infeasible, regardless of other changes to the 0–1 integer program. By considering the set of all partial solutions attributed to k, (which he called Λ_k), one can obtain the minimum relaxation in k before *any* partial solution in Λ_k can become potentially feasible. Moreover, only partial solutions in Λ_k need be examined if one relaxes constraint k, since partial solutions fathomed by other constraints remain fathomed.

Thus, Roodman laid the foundation for the work presented here. Our contributions may be summarized as follows:

1) The sets Λ_k $(k=0, 1, \ldots, m)$ have been purged of all fathomed partial solutions attributable to more than a single constraint, thus improving the ranging formulas.

2) A method is outlined for dealing with problem tightening. (Roodman dealt only with problem relaxation.)

---

*This research was performed as part of the activities of the Management Sciences Research Group, Carnegie-Mellon University when the authors were at that institution.

3) An efficient storage scheme is presented in recognition of the magnitude of the data collection and classification task implicit in this approach to postoptimization.

4) A unified set of detailed algorithms is presented that use the storage structure to postoptimize and re-postoptimize after one or more problem parameter changes. Thus, a decision maker can change several parameters, analyze the results, and then change several more parameters efficiently.

5) The difficulties and trade-offs inherent in the use of surrogate constraints and imbedded linear programs are discussed from an implementation view-point. Appropriate ranging formulas and classification rules are also provided.

This paper has four sections. Section I presents an implicit enumeration algorithm for solving 0–1 integer programs and classifying all fathomed partial solutions generated in the process. Section II presents formulas and algorithms that can be respectively used for parameter ranging, and for locating those partial solutions which need to be examined in light of single changes in problem parameters. Section III discusses multiparameter postoptimization and its effects upon the collected data. Also contained in Section III is an analysis of the effects of adding surrogate constraints and/or imbedded linear programs to the implicit enumeration algorithm. Finally, the paper is concluded with a numerical example in Section IV.

I. AN IMPLICIT ENUMERATION ALGORITHM FOR CLASSIFYING FATHOMED PARTIAL SOLUTIONS

This section presents an implicit enumeration algorithm which gathers the information required for postoptimizing a 0–1 integer program.

Consider the integer program, IP, defined as follows:

$$\min \sum_{j=1}^{n} c_j x_j \tag{1}$$

$$\text{s.t.} \sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad \text{for } i \in M = \{0, 1, \ldots, m\} \tag{2}$$

$$x_j = 0 \text{ or } 1 \quad \text{for } j \in N = \{1, 2, \ldots, n\} \tag{3}$$

For expository purposes we shall assume $c_j \geq 0$ for all $j \in N$. At the cost of additional algorithmic complexity, this assumption can be readily dropped. Define $a_{0j} = c_j$ $(j \in N)$ and $b_0 = \bar{z}$ where $\bar{z}$ is a valid upper bound on the optimal solution. Row zero is the "objective function constraint" and is to be continually updated during the search.

If you take only (3) into consideration, there are 2^n possible assignments of values to $(x_1, x_2, \ldots, x_n)$. Let an assignment of values to a subset of these variables be called a partial solution. Let the variables x_j not assigned values be called free variables. Any specific assignment of values to the free variables is called a completion of the associated partial solution. If a partial solution contains d variables then there are 2^{n-d} completions. The solution of IP associated with a partial solution is the completion whose free variables are all zero. Call this the zero-completion.

As the implicit enumeration search proceeds, partial solutions are generated in an attempt to find a feasible zero-completion. Nonzero-completions of a feasible partial solution cannot have a better objective function value since $c_j \geq 0$ for all $j \epsilon N$. A partial solution is fathomed (a) if its zero-completion is feasible, or (b) if it can be shown that none of its completions can yield a feasible solution better than the best feasible solution found to date. The 2^{n-d} completions of a fathomed solution need not be considered.

After finding each feasible zero-completion, the objective function constraint is updated so that only solutions having lower or equal objective function values are considered. For example, suppose the best feasible zero-completion found to date has objective function value $\tilde{z}$, then the objective function constraint becomes $\sum_{j=1}^{n} a_{0j}x_j \leq \tilde{z}$. That is, b_0 is set equal to $\tilde{z}$.

Each fathomed partial solution will be classified in two ways:

(i) by the identity of the fathoming constraint(s); and

(ii) by the value of its objective function.

The data structure to be described in the second half of this section permits a reasonably straightforward implementation of this dual classification. To obtain a convenient indexing procedure for the objective function values of fathomed partial solutions, define an ordered set of scalars $\{\tilde{z}_0, \ldots, \tilde{z}_\tau\}$, where $\tilde{z}_0 < \ldots < \tilde{z}_\tau$, and where $\tilde{z}_0$ and $\tilde{z}_\tau$ are respectively lower and upper bounds on the objective function values. Possible values for $\tilde{z}_0$ and $\tilde{z}_\tau$ are 0 and Σc_j. Define, furthermore, a function, H, over this range of objective values, such that if some $\hat{z}$ is contained in the interval $\tilde{z}_{u-1} \leq \hat{z} < \tilde{z}_u$, then $H(\hat{z}) = u$. For example, if $\tilde{z}_u - \tilde{z}_{u-1} = 1$ for $u = 1, \ldots, \tau$ and $\tilde{z}_0 = 0$, then $H(\hat{z}) = \hat{z}$. Thus, fathomed partial solutions with objective function value $\hat{z}$ can be indexed by the identity of the fathoming constraint(s) and by $H(\hat{z})$.

For partial solution s define $x(s)$ as the solution of IP associated with s (i.e. the zero-completion) and $z(s)$ as the value of its objective function. Let $\{s^1, \ldots, s^p\}$ be the set of fathomed partial solutions with feasible zero-completions generated during the algorithm. Let $x^1 = x(s^1), \ldots, x^p = x(s^p)$ be the associated solutions with objective function values $z^1 = z(s^1), \ldots, z^p = z(s^p)$. Set $z^0 = \tilde{z}_\tau$. Assume $z^0 \geq z^1 \geq \ldots \geq z^p$.

Define the following sets of fathomed partial solutions for $u = 1, \ldots, \tau$:

(4) $\Lambda_k^u = \{s \mid$ the zero-completion of s is infeasible and constraint k is the *only* constraint that fathoms s; and $u = H(z(s))\}$; for $k = 0, 1, \ldots, m$.

(5) $\Lambda_g^u = \{s \mid$ the zero-completion of s is infeasible and s has two or more fathoming constraints; and $u = H(z(s))\}$

(6) $\Lambda_f = \{s^1, \ldots, s^p\} = \bigcup_{u=1}^{\tau} \Lambda_f^u$ where $s \in \Lambda_f^u$ if the zero-completion of s is feasible and $u = H(z(s))$.

$$\Lambda_k = \bigcup_{u=1}^{\tau} \Lambda_k^u \qquad \text{for } k = 0, 1, \ldots, m$$

$$\Lambda = \Lambda_f \cup \Lambda_g \cup \bigcup_{k=0}^{m} \Lambda_k$$

The sets Λ_r^u ($r=f$, g, 0, 1, . . ., m; $u=1$, . . ., τ) form a partition of all fathomed partial solutions generated during the solution of IP using implicit enumeration. Section II shows how these sets are used to determine ranges of the problem parameters for which the optimal solution to IP remains optimal, and also shows how these sets are used to postoptimize IP if some parameters exceed the range.

These definitions differ from Roodman's [9] definition of Λ_k ($k=0$, 1, ... m). An example of how these definitions help to give better ranges for the problem parameters than those achieved by Roodman is given in Section II. Also, see the numerical example in Section IV.

Associate the following index sets with each partial solution obtained during the search:

$$J_s^1 \doteq \{j \epsilon N \mid x_j(s)=1\}; \qquad J_s^0 \doteq \{j \epsilon N \mid x_j(s)=0\}; \qquad \text{and } N_s \doteq N - J_s^1 \doteq J_s^0,$$

where $x_j(s)=d$ implies that the jth variable is assigned value d ($d=0$ or 1) in the partial solution s. J_s^1 and J_s^0 are index sets for the fixed variables, and N_s is the index set of free variables with respect to s.

Using these index sets, define the following variables for each row $i \epsilon M$:

$$t_i(s) = b_i - \sum_{j \epsilon J_s^1} a_{ij}; \tag{7}$$

$$\rho_i(s) = \sum_{j \epsilon N_s} \min(a_{ij}, 0); \text{ and} \tag{8}$$

$$\gamma_i(s) = t_i(s) - \rho_i(s). \tag{9}$$

The latter quantity, $\gamma_i(s)$, can be interpreted as the upper bound on the slack in constraint i, that can result from any completion of s. Consequently, negative values of $\gamma_i(s)$ are not permitted to occur during the algorithm, since this would imply the infeasibility of all completions of s. The infeasibility test described below, when applied as indicated in Algorithm 1 will ensure that $\gamma_i(s)$ remains nonnegative—providing it is nonnegative for the initial (all-zero) solution.

The algorithm tests for the following two implications:

Feasibility: If $t_i(s) \geq 0$ for all $i \epsilon M$, partial solution s is feasible.

Infeasibility: For any partial solution s, and $j \epsilon N_s$, define the following pair of index sets.

$$G^1 = \{i \epsilon M \mid \gamma_i(s) < -a_{ij}\}$$
$$G^0 = \{i \epsilon M \mid \gamma_i(s) < a_{ij}\} \tag{10}$$

Then, it can readily be shown that $G^d \neq \phi$ ($d=1$ or 0), is a sufficient condition for all completions of s with $x_j(s)=1-d$ to be infeasible.

The algorithm to be presented differs slightly from the usual implicit enumeration scheme in the way its infeasibility test or "skipping step" is arranged. It is conventional practice to test all variables $j \epsilon N_s$ against a single constraint, i, thus forming the following index sets.

$$F_i^1 = \{j \epsilon N_s \mid \gamma_i(s) < -a_{ij}\};$$

and

$$F_i^0 = \{j \in N_s \mid \gamma_i(s) < a_{ij}\} \qquad \text{for each } i \in M.$$

These sets are then combined to form $F^d = \bigcup_{i \in M} F_i^d$ ($d=1$ and 0). Now, if $F^0 \cap F^1$ is nonempty, then a contradiction exists, and all completions of s are infeasible. Otherwise, variables contained in F^1 and F^0 (if any) may be fixed (skipped) at 1 and 0 respectively.

Although the sets F_i^1, F_i^0 ($i=0, \ldots, m$) contain the required classification information, they must be rearranged so that each fathomed partial solution is attributed to each of its fathoming constraints. This is not the case with the type of infeasibility test proposed above. Here, variables $j \in N_s$ are tested one at a time against all constraints to form the sets G^1 and G^0 defined in (10). If G^d ($d=1$ or 0) is nonempty, then the contents of G^d can be used to update Λ_r directly. A second benefit of the approach comes from the fact that the infeasibility test becomes progressively stronger, since $\gamma_i(s)$, $i \in M$, is updated (in a nonincreasing sequence) each time a variable is skipped.

Algorithm 1, which follows immediately, may be used to classify and collect partial solutions while solving IP. Certain details are omitted. No provision is made for the use of imbedded linear programs, or the use of surrogate constraints. This will be discussed in Section III. Termination criteria, rules for generating new partial solutions, and rules for backtracking from old ones, are not given. Any of the methods proposed in [1, 3, 4, 5, 6] can be used, since the classification and collection procedures are independent of them.

Algorithm 1

1. Obtain initial values of p, J_s^1, J_s^0, N_s, and Λ_r ($r=f, g, 0, 1, \ldots, m$). [When Algorithm 1 is used to solve IP, initialize $p=0$, $J_s^1=\emptyset$, $J_s^0=\emptyset$, $N_s=N$, $\Lambda_r=\emptyset$ for $r=f, g, 0, 1, \ldots, m$. Subsequently, when Algorithm 1 is used, during postoptimization, to examine the completions of partial solution s then J_s^1, J_s^0, N_s describe s; and p and Λ_r describe the current state of the postoptimization.]
2. Set $\alpha=n+1$, $j=0$, and go to 3.
3. Compute $t_i(s)$, $\rho_i(s)$, and $\gamma_i(s)$ using (7)–(9).
   If $t_i(s) < 0$ for some $i \in M$, go to 4.
   Otherwise, s is feasible with objective function value $z(s)$.
   Add s to Λ_f.
   Set $b_0=\bar{z}=z(s)$, and $p=p+1$.
   Backtrack, and go to 2.
4. Set $j=j+1$.
   If $j=\alpha$, *Generate*, and go to 2.
   If $j>n$, set $j=0$ and repeat 4.
   If $j \notin N_s$, repeat 4.
   Otherwise, go to 5.
5. Compute G^1 and G^0 using (10).
   If $G^1=\emptyset$ and $G^0=\emptyset$, go to 4.
   If $G^d \neq \emptyset$ and $G^{1-d}=\emptyset$ ($d=1$ or 0), go to 6.
   Otherwise, there is no feasible completion of s.

For $d=1$ and 0:

If $G^d=\{k\}$, add $\bar{s}$ to Λ_k^h, where
$J_{\bar{s}}^d=J_s^d$, $J_{\bar{s}}^{1-d}=J_s^{1-d}\cup\{j\}$, and
$h=H(z(\bar{s}))$.

Otherwise, add $\bar{s}$ (as defined) to Λ_g^h.

Backtrack, and go to 2.

6. If $G^d=\{k\}$, add $\bar{s}$ to Λ_k^h, where
$J_{\bar{s}}^d=J_s^d$, $J_{\bar{s}}^{1-d}=J_s^{1-d}\cup\{j\}$, and
$h=H(z(\bar{s}))$.
Otherwise, add $\bar{s}$ (as defined) to Λ_g^h.
Skip to partial solution $\tilde{s}$, where
$J_{\tilde{s}}^d=J_s^d\cup\{j\}$, $J_{\tilde{s}}^{1-d}=J_s^{1-d}$.
Set $s=\tilde{s}$, $\alpha=j$, and go to 3.

The practicality of the approach to postoptimizing zero-one programs, that is being described, hinges on the ability of the computing system to conveniently store and access the subsets that compose Λ. Although the specific details of the storage mechanisms are design considerations, that are best deferred until implementation, the following discussion is included to demonstrate the plausibility of our procedures.

Of all the likely data structures, the linked list seems the most appropriate storage scheme. See Knuth [8] for a discussion of this and other possible information structures. Consider a list composed of "records" that contain the following information for $s\in\Lambda_r^u$ $(r=f, 0, 1, \ldots, m)$:

1) the fathomed partial solution, s;
2) the value of $z(s)$.
3) the value of $\gamma(s)$.
4) a pointer, $e(s)$, indicating the location in the list of the previous fathomed partial solution attributable to this set; and
5) a pointer, $f(s)$, indicating the location of the next fathomed partial solution belonging to this list or to Λ_g^u (with constraint r among the fathoming constraints), whichever is fathomed first.

From this description, one can see that a modified, doubly linked list is proposed. Although storage for $s'\in\Lambda_g$ is not provided in this list, the locations in Λ_g of such s' are retained via pointer $f(s)$. Thus, these partial solutions can be accessed, updated and possibly reclassified after changes are made to problem parameters. Details on storing Λ_g will be omitted, since they parallel those for Λ_r. It suffices to note that each record of Λ_g must contain the relevant information on each fathoming constraint.

The list would be referenced through a "list head." This core-resident array, L, would have an element allocated for each constraint and each objective function interval. The value of any one of these $m\tau$ elements indicates the location of the last member of the set for which it is the head. Consequently, L would be updated each time a partial solution is added to Λ. Any of the sets composing Λ, can be retrieved, therefore, by beginning with the partial solution indicated by the list head, and following the pointers until all set elements are obtained.

The sets J_s^1, J_s^0, and N_s, that describe partial solution s, can be most compactly stored by using a single bit to indicate whether a variable is free or fixed, followed by a second bit to indicate whether the

variable, if fixed, is zero or one. A small saving in storage can be achieved, at some added computation cost, by omitting this second bit if the variable is free.

Without this economy measure, the following memory requirements would exist for each partial solution, s.

ENTITY	STORAGE REQUIREMENT
s	2 bits per variable
$z(s)$	1 word
$\gamma_k(s)$	1 word
$e(s)$	1 word
$f(s)$	1 word

Problems in the range of 30 variables would thus require about six words of storage per partial solution. This estimate is based on the assumption that full words (32–36 bits) are used throughout. Further memory savings can be obtained by using half-words, or bit packing.

Records so constructed would be added to the list sequentially. A temporary core-resident buffer of several thousand words (or larger, if possible), would hold the list. When full, the buffer would be emptied into secondary storage (i.e. tape or disc), and reused.

II. SINGLE PARAMETER POSTOPTIMIZATION

In Section I, an algorithm was presented which solved IP and gathered the set of fathomed partial solutions, Λ (i.e. $\sum_{k=0}^{m} \Lambda_k \cup \Lambda_f \cup \Lambda_g$). Different generate procedures, and/or surrogate constraints will create different Λ sets. (See Section III for a discussion of the classification procedures needed when partial solutions are fathomed by surrogate constraints.)

In this section, formulas will be presented which provide ranges on the amount single parameters may be altered before some fathomed partial solution, $s \in \Lambda$, ceases to be fathomed. Completions of such an s may yield new optima. Efficient algorithms are also presented for locating those s in Λ, whose completions should be examined, in order to reoptimize IP following any single, parameter change. The algorithms indicate the sequence in which the completions of these partial solutions should be examined by Algorithm 1. Partial solution s should be examined before partial solution $\bar{s}$ if $z(s) < z(\bar{s})$, since s is more likely to yield the optimum than $\bar{s}$. Further, s may have a feasible descendant, say $\bar{s}$, with $z(\bar{s}) < z(\bar{s})$ in which case $\bar{s}$ need not be examined at all.

The ranging formulas to be presented, that pertain to problem relaxation, depend on the composition of Λ. Thus, these ranges are not ranges in the usual linear programming sense, since it is quite possible that parameter alterations to values outside the range will not change the optimal solution. For this reason, we suspect that a user may find the algorithms more helpful than the ranging formulas.

Postoptimization procedures will be presented for changes in the right hand sides, in the objective function coefficients, and in the matrix elements. In each case, parameters may be altered so as to "tighten" or "relax" the current problem.

Ranging formulas arrived at by relaxing the problem (i.e. increasing the right hand sides, and

decreasing the objective coefficients or matrix elements) are essentially those given by Roodman. In these cases, the primary difference relates to our removal of certain partial solutions from Roodman's Λ_k sets if more than one constraint causes fathoming. The favorable consequences of this change are illustrated in the example at the end of this section.

The procedures related to tightening the problem are our own. Their existence eliminates the need for running the optimization phase with the tightest possible problem, as proposed by Roodman. Thus, the user may commence optimization with the problem defined by the set of parameter values of most interest to him. The ability to vary a parameter in both directions provides more flexibility during postoptimization—particularly when analyzing multiple parameter alterations. This will be discussed in Section III.

Three additional index sets will now be defined that simplify the notation used in stating the ranging formulas. Let $x^q, \ldots, x^p$ be the optimal solutions to IP. Then define the index sets as follows:

$$\Lambda_{f'} = \{s^q, \ldots, s^p\};$$

and

$$Q_r^d = \{s \in \Lambda_{f'} \mid x_r(s) = d\}, \qquad d = 0, 1.$$

[Recall that $x_r(s)$ is the rth component of the zero-completion of s.]

Altering the Right Hand Sides

Let b_k^- and b_k^+ be the lower and upper bounds on the amount by which b_k can be changed before some fathomed partial solution ceases to be fathomed. [That is, for changes in b_k to b_k', where $b_k - b_k^- \leq b_k' \leq b_k + b_k^+$, all fathomed partial solutions remain fathomed.] These values may be determined as follows.

Define for each $s \in \Lambda_k$, $k \neq 0$ $\left(\text{recall } \Lambda_k = \bigcup_{u=1}^{\tau} \Lambda_k^u\right)$

$$b_k^+(s) = -1 - \gamma_k(s),$$

then

$$b_k^+ = \min_{s \in \Lambda_k} \{b_k^+(s) \mid z(s) \leq z^p\}$$

$$= \min_{s \in \bar{\Lambda}_k} \{b_k^+(s)\} \qquad \text{where } \Lambda_k = \bigcup_{u=0} \Lambda_k^u \qquad \text{and } u' = H(z^p).$$

Similarly, define for each $s \in \Lambda_{f'}$

$$b_k^-(s) = t_k(s),$$

then

$$b_k^- = \max_{s \in \Lambda_{f'}} \{b_k^-(s)\}.$$

The value of b_k^- is simply taken from the optimal solution with the largest slack in row k. This solution will be feasible for all decreases in b_k less than or equal to b_k^-. The upper bound, b_k^+, is taken from the infeasible fathomed partial solution, s, with the smallest potential infeasibility, $-\gamma_k(s)$, in the kth row.

The following two algorithms indicate which partial solutions are no longer fathomed after changes in b_k outside the above range. Hence, the algorithms point to the partial solutions which need to be examined in order to find new optimal solutions (if they exist).

Algorithm 2 supervises the enumeration of the completions of partial solutions in Λ_k which may have feasible completions with objective function values better than the current optima in view of an increase in b_k. These completions are classified by Algorithm 1. Here advantage is taken of the fact that Λ_k is indexed by objective function value by examining the completions of such partial solutions in order of ascending objective function value. Hence partial solutions in Λ_k^w are examined before those in Λ_k^v if $w < v$. The information maintained for each fathomed partial solution in Λ_k is updated to reflect the change in b_k. When Algorithm 2 is finished, postoptimization due to another parameter change can commence.

Algorithm 2: Increase in b_k to b_k'

(Better solutions than the current optima may exist if $b_k' > b_k + b_k^+$.)

1. Set $\bar{x} = x^p$, $\bar{z} = z^p$, $v = 0$. Set $\bar{\Lambda} = \Lambda_k^v$. The current objective constraint is $cx \leqslant \bar{z}$. Go to 2.
2. If $\bar{\Lambda} = \emptyset$, go to 5.
   Otherwise, take any $s \in \bar{\Lambda}$. Set $\bar{\Lambda} = \bar{\Lambda} - \{s\}$. Go to 3.
3. If $b_k' > b_k + b_k^+(s)$ go to 4.
   Otherwise, recompute $\gamma_k(s)$ to reflect the change in b_k and go to 2.
4. Set $\Lambda_k^v = \Lambda_k^v - \{s\}$ and examine the completions of s using Algorithm 1. If all completions are infeasible, go to 2. Otherwise, let $\bar{s}$ be the best feasible completion. [Note that the objective function constraint is considered as part of the constraint set.] Set $\bar{x} = x(\bar{s})$, $b_0 = \bar{z} = z(\bar{s})$, update the objective constraint to $cx \leqslant \bar{z}$, and go to 2.
5. Set $v = v + 1$. If $\bar{z} \geqslant \tilde{z}_{v-1}$, set $\bar{\Lambda} = \Lambda_k^v$ and go to 2.
   Otherwise stop: $\bar{x}$ is the optimal solution to the new problem.
   Also, update $\gamma_k(s)$ for $s \in \Lambda_k^1 \cup \ldots \cup \Lambda_k^v$ to reflect the changed parameters.

Algorithm 3 points to the partial solutions which need to be examined if b_k is decreased. The algorithm first checks those partial solutions which had a feasible zero completion prior to the change. If one of the previous optima is still feasible then it remains optimal. If none are feasible, then the best feasible partial solution in Λ_f (if one exists) will provide an initial objective value bound used in the search for the new optimum. Next, completions of those solutions in Λ_f found infeasible after decreasing b_k and having objective value smaller than the initial objective value bound are examined and classified. Finally, completions of partial solutions, that were fathomed because it was ascertained that they could not yield better solutions than the best found to date, are examined and classified, (i.e. Λ_0 is examined). After completions of all previously fathomed partial solutions, which are no longer fathomed after the change in b_k, are classified by Algorithm 1, further postoptimization can begin.

Algorithm 3: Decrease in b_k to b_k'

(Optimal solutions are no longer feasible if $b_k' < b_k - b_k^-$.)

1. Update $\gamma_k(s)$ for $s \in \Lambda_k$ to reflect the changed parameter. Set $t = p$, $b_0 = \bar{z} = \tilde{z}_t$. Go to 2.

2. If $t=0$, none of the previously feasible solutions is feasible for the new problem. Set $t=p$. Go to 5.
   Otherwise, go to 3.
3. If x^t is infeasible, reassign s^t to Λ_0, set $t=t-1$, and go to 2. Otherwise, x^t is the current best solution. Set $\bar{x}=x^t$, $\bar{z}=z^t$.
   The objective constraint is $cx \leqslant \bar{z}$. Set $t=t-1$.
   Go to 4.
4. If $t=0$, set $t=p$, and go to 5.
   If x^t is infeasible, reassign s^t to Λ_0.
   Set $t=t-1$ and repeat 4.
5. If $\bar{z} \geqslant z^t$, go to 6.
   Otherwise, set $v=0$, $\bar{\Lambda}=\Lambda_0^v$, and go to 7.
6. Examine the completions of s^t using Algorithm 1. If all completions are infeasible, set $t=t-1$, and go to 5. Otherwise, let $\bar{s}$ be the best feasible completion. Set $\bar{x}=x(\bar{s})$, $b_0=\bar{z}=z(\bar{s})$, $t=t-1$, update the objective constraint to $cx \leqslant \bar{z}$ and go to 5.
7. If $\bar{\Lambda}=0$, go to 9.
   Otherwise, take any $s \in \bar{\Lambda}$. Set $\bar{\Lambda}=\bar{\Lambda}-\{s\}$.
   Go to 8.
8. Set $\Lambda_0^v=\Lambda_0^v-\{s\}$. Examine the completions of s using Algorithm 1. If all completions are infeasible, go to 7.
   Otherwise, let $\bar{s}$ be the best feasible completion. Set $\bar{x}=x(\bar{s})$, $b_0=\bar{z}=z(\bar{s})$, update the objective constraint to $cx \leqslant \bar{z}$, and go to 7.
9. Set $v=v+1$. If $\bar{z} \geqslant \tilde{z}_{v-1}$, set $\bar{\Lambda}=\Lambda_0^v$ and go to 7.
   Otherwise, stop: If $\bar{z}=\infty$, no feasible solution exists.
   If $\bar{z}<\infty$, $\bar{x}$ is the optimal solution to the new problem.

Altering the Objective Function Coefficients

Let c_r^- and c_r^+ be the lower and upper bounds on the amount by which $c_r \geqslant 0$ can be changed before some fathomed partial solution ceases to be fathomed. [That is, for changes in c_r to c'_r where $c_r - c_r^- \leqslant c'_r \leqslant c_r + c_r^+$.] These values may be determined as follows:

For each $s \in \Lambda_0 \cup \Lambda_f - \Lambda_{f'}$

define

$$c_r^-(s) = \begin{cases} \min\{z(s)-z^p,\ c_r\} & \text{if } Q_r^1=\emptyset \text{ and } r \in J_s^1 \\ c_r & \text{otherwise} \end{cases}$$

$$c_r^+(s) = \begin{cases} z(s)-z^p & \text{if } Q_r^0=\emptyset \text{ and } r \notin J_s^1 \\ \infty & \text{otherwise.} \end{cases}$$

Then

$$c_r^- = \min_{s \in \Lambda_0 \cup \Lambda_f - \Lambda_{f'}} \{c_r^-(s)\},$$

and

$$c_r^+ = \min_{s \in \Lambda_0 \cup \Lambda_f - \Lambda_{f'}} \{c_r^+(s)\}.$$

Ranges are thus determined by the smallest change in c_r that may cause the objective function value of any partial solution belonging to Λ_f or to Λ_0 to decrease below the value of the best solution in $\Lambda_{f'}$ after changing c_r.

Changes in c_r beyond these values may result in new optimal solutions. Algorithm 4 is used if c_r is decreased or increased. The algorithm first updates $z(s)$ for all partial solutions, s, in Λ. Next, the best solution in Λ_f is determined. Finally, the completions of those partial solutions in Λ_0 which may yield a still better solution are examined.

Algorithm 4: Decrease or Increase c_r to $c_r' \geq 0$

(New optima may occur if $c_r' < c_r - c_r^-$ or $c_r' > c_r + c_r^+$.)

1. Update $z(s)$ for $s \in \Lambda$ if $r \in J_s^1$. Rearrange the memberships in the various linked lists so that $s \in \Lambda_r^v$ if and only if $H(z(s)) = v$, $(r = f, g, 0, \ldots, m)$. Let $\bar{x} = x(\bar{s})$ be the solution in Λ_f with the lowest (best) objective value. Set $b_0 = \bar{z} = z(\bar{s})$. Let $v = 0$ and $\bar{\Lambda} = \Lambda_0^v$.
2. If $\bar{\Lambda} = \phi$. Go to 4.
   Otherwise, take any $s \in \bar{\Lambda}$. Set $\bar{\Lambda} = \bar{\Lambda} - \{s\}$. Go to 3.
3. If $z(s) > \bar{z}$, go to 2.
   Otherwise, set $\Lambda_0^v = \Lambda_0^v - \{s\}$ and examine the completions of s using Algorithm 1. Go to 2.
4. Set $v = v + 1$. If $\bar{z} \geq \tilde{z}_{v-1}$, set $\bar{\Lambda} = \Lambda_0^v$ and go to 2.
   Otherwise, stop: $\bar{x}$ is the optimal solution to the new problem.

Altering the Matrix Elements

To obtain a lower bound, $a_{\overline{kr}}$, on the extent any given matrix element, a_{kr}, can be decreased before the current optima may cease to be so, define for $s \in \Lambda_k$, $k \neq 0$

$$a_{\overline{kr}}^-(s) = \begin{cases} -1 - \gamma_k(s), & \text{if } r \in J_s^1 \\ -1 - \gamma_k(s), & \text{if } r \in N_s \text{ and } a_{kr} < 0 \\ -1 - \gamma_k(s) + a_{kr} & \text{if } r \in N_s \text{ and } a_k \geq 0 \\ \infty & \text{if } r \in J_s^0 \end{cases}$$

Then

$$a_{\overline{kr}}^- = \min_{s \in \Lambda_k} \{a_{\overline{kr}}^-(s) \mid z(s) \leq z^v\}$$

$$= \min_{s \in \tilde{\Lambda}_k} \{a_{\overline{kr}}^-(s)\}, \qquad \text{where } \tilde{\Lambda}_k = \bigcup_{u=0}^{u'} \Lambda_k^u, \qquad \text{and } u' = H(z^v).$$

The justification for this result follows the reasoning used in obtaining b_k^+. Note that if $a_{kr} \leq \gamma_k(s)$, then constraint k ceases to fathom s. Hence, if variable r is free in s, then the value of nonnegative coefficient, a_{kr}, must be added to obtain $a_{\overline{kr}}^-(s)$.

The upper bound, $a_{\overline{kr}}^+$, may be found analogously by b_k^- by first defining for $s \in \Lambda_{f'}$, and $k \neq 0$:

$$a_{\overline{kr}}^+(s) = \begin{cases} t_k(s) & \text{if } Q_r^0 = \emptyset \\ \infty & \text{otherwise.} \end{cases}$$

Then,

$$a_{kr}^{+} = \max_{s \in \mathcal{N}} \{a_{kr}^{+}(s)\}.$$

Changes in a_{kr} beyond a_{kr}^{+} will result in new solutions being obtained. If a_{kr} is changed to a'_{kr}, where $a_{kr} - a_{kr}^{-} \leq a'_{kr} \leq a_{kr} + a_{kr}^{+}$, then all fathomed partial solutions remain fathomed. For changes outside this range, Algorithms 5 and 6 are applicable.

Algorithm 5: Decrease a_{kr} to $a'_{kr} < a_{kr} - a_{kr}^{-}$

(Better solutions than the current optima may exist.)

This is the same as Algorithm 2 except for the following replacement for step 3.

3. If $a'_{kr} < a_{kr} - a_{kr}^{-}(s)$ go to 4.
   Otherwise, recompute $\gamma_k(s)$ to reflect the change in a_{kr} and go to 2.

Algorithm 6: Increase a_{kr} to $a'_{kr} > a_{kr} + a_{kr}^{+}$

(Optimal solutions are no longer feasible.)

This is the same as Algorithm 3 except for the following replacement for step 7.

7. If $\bar{\Lambda} = \emptyset$, go to 9.
   Otherwise, take any $s \in \bar{\Lambda}$. Set $\bar{\Lambda} = \bar{\Lambda} - \{s\}$.
   If $r \notin J_s^1$, go to 8; otherwise, repeat 7.

Algorithms 2 through 6 are of three types. If the problem parameters are changed so as to tighten the problem constraints to a point at which the optimal solution(s) is(are) no longer feasible, then the search for a new optimum strives toward finding the best feasible solution with the new constraints. Previously infeasible fathomed partial solutions need not be examined, since they remain infeasible. If a new optimum is found, it will have objective function value higher than the old optimum. Tightening the problem constraints is considered in Algorithms 3 and 6.

On the other hand, if the problem parameters cause the constraints to be relaxed, the old optima remain feasible. However, there is a possibility that another, previously infeasible solution, may become feasible and yield a better objective function value. Algorithms 2 and 5 show where to look for such a solution.

If the costs, $c_j (j = 1, \ldots, n)$, are changed, then feasible solutions remain feasible, but the optima may not remain optimal. If new optimal solutions are to found, Algorithm 4 shows where to look.

As mentioned previously, the Λ_k sets defined in this paper (definitions (4)–(6)) are different than those defined in [9]. If $\bar{\Lambda}_k$ denotes the Λ_k set defined by Roodman, then $\Lambda_k \subseteq \bar{\Lambda}_k$. The reason for this is that Roodman does not define a Λ_g set. Instead, all partial solutions are lumped into some Λ_k set for $k = 0, 1, \ldots, m$. Since $\Lambda_k \subseteq \bar{\Lambda}_k$, $\bar{b}_k^{+} \leq b_k^{+}$ where $\bar{b}_k^{+} = \min_{s \in \bar{\Lambda}_k} \{b_k^{+}(s)\}$. Similarly $\bar{c}_r^{-} \leq c_r^{-}$ and $\bar{a}_{kr}^{-} \leq a_{kr}^{-}$, where $\bar{c}_r^{-}$ and $\bar{a}_{kr}^{-}$ are defined in terms of $\bar{\Lambda}_k$. Hence, the ranges for parameters arrived at using definitions (4)–(6) are at worst the same as the ranges defined by Roodman. To illustrate the potential difference, consider the following constraint set.

$$x_1+x_2 \leq 0$$
$$x_3+x_4 \leq 0$$
$$x_1+x_2+x_3+x_3 \leq 0$$

Then the fathomed partial solutions are $s_1=(1, 0, 0, 0)$, $s_2=(0, 1, 0, 0)$, $s_3=(0, 0, 1, 0)$ and $s_4=(0, 0, 0, 1)$. These partial solutions are partitioned as follows:

$$\bar{\Lambda}_1=\{s_1, s_2\},\ \bar{\Lambda}_2=\{s_3, s_4\},\ \bar{\Lambda}_3=\emptyset; \qquad \Lambda_1=\emptyset,\ \Lambda_2=\emptyset,\ \Lambda_3=0,\ \Lambda_g=\{s_1, s_2, s_3, s_4\}.$$

Hence

$$\bar{b}_1^+=1,\ \mathbf{b}_2^+=1,\ \bar{b}_3^+=\infty; \qquad b_1^+=\infty,\ b_2^+=\infty,\ b_3^+=\infty.$$

The numerical example in section IV is the same as the one in [9] and it, too, shows that better ranges can be achieved via definitions (4)–(6).

III. ADDITIONAL REMARKS

Multiple Parameter Postoptimization

Changes in two or more parameters can be readily accommodated by the algorithms already presented. The key to this ability is the continual modification (and enlargement) of Λ that occurs during postoptimization. If a parameter change is sufficient to cease the fathoming of a partial solution, s, then s is replaced in Λ by the set of fathomed partial solutions created by examining the descendents of s. In each case where an examination of a partial solution is prescribed, Algorithm 1 is employed. Thus, Λ gradually expands as parameters are adjusted up and down. It should be pointed out that the set of solutions to IP associated with the partial solutions in Λ also increases and that every solution to IP is a descendent of one of these solutions. Ultimately, all 2^n unique arrangements could be explicitly obtained and classified if the procedure is pursued long enough. Hence, a sequence of single parameter postoptimization runs can be used to effect any desired set of multiple parameter changes.

In comparing our procedures with Roodman's, one finds that the relative utility of Roodman's "temporary" parameter adjustment decreases considerably because of our ability to tighten, as well as relax parameter settings. A temporary change from x to y can thus be replaced with a permanent change from x to y, followed by a second, permanent change from y back to x. Clearly, no new completions of partial solutions need be explored in reverse changes, since the earlier optimum is already explicitly contained in Λ. This combination of two permanent changes also possesses the desirable property of enlarging Λ.

Up until this point, the maintenance of Λ_g has not been discussed. It is important to note, however, that the members of Λ_g must be examined, and the associated information updated, after each parameter change. Whenever the set of fathoming constraints of $s\epsilon\Lambda_g$ diminishes to one constraint, say k, s must be removed from Λ_g, and placed in Λ_k.

Accomodation of Surrogate Constraints and Imbedded Linear Programs

For any partial solution, s, consider forming a linear combination of the constraints $i\epsilon M$ using weights $u_i \geq 0$. Let $M'=\{i\epsilon M | u_i>0\}$. Then, the surrogate, or composite, constraint can be defined, as constraint $m+1$, as follows:

$$\sum_{j\in N_s}\left(\sum_{i\in M'} u_i a_{ij}\right)x_j \leq \sum_{i\in M'} u_i\left(b_i - \sum_{j\in J_s^1} a_{ij}\right) \tag{11}$$

It has been shown in practice that the use of surrogate constraints increases the efficacy of implicit enumeration algorithms. Two reasons may be advanced:

1) The surrogate is usually "different" from any one of the original constraints, and hence, may provide additional information when the various implicit enumeration tests are applied; and
2) Since the surrogate is obtained from the original constraints, and therefore, somehow "represents" them, it is sometimes chosen for use with additional powerful, but time-consuming, tests that cannot be efficiently applied to each of the original constraints individually.

Balas [2], Geoffrion [5], and Glover [7] have proposed criteria for defining a "strongest" surrogate constraint.

Effective use of surrogate constraints causes earlier fathoming of partial solutions. Hence, the Λ set formed in Algorithm 1 becomes smaller. From the stand-point of postoptimization, however, classification of any partial solution, s, fathomed by the surrogate becomes less precise. If a surrogate had not been used, the completions of s would have been examined until each had been fathomed by one or more constraints, $k\epsilon M$. Since fathoming now occurs before these partial solutions and fathoming constraints are identified, s must be added to one or more Λ_k sets. Classification of s follows three cases:

C1) If the surrogate is the *only* constraint which fathoms s; then add s to Λ_k for each $k\epsilon M'$.
C2) If the surrogate and one constraint k, for $k\epsilon M'$, fathom s; then add s only to Λ_k.
C3) If the surrogate and constraint k, for $k\notin M'$, fathom s; or, if the surrogate and two or more constraints fathom s; then add s to Λ_g.

Ranging for partial solutions fathomed as in cases C1 and C2 can be accomplished. Let

$$w_k = -1 - \frac{\gamma_{m+1}(s)}{u_k}, \qquad \gamma_{kr} = \frac{a_{m+1,r}}{u_k}.$$

Then for:

C1) $b_k^+(s) = w_k$; and

$$a_{kr}^-(s) = \begin{cases} w_k & \text{if } r\epsilon J_s^1 \\ w_k & \text{if } r\epsilon N_s \text{ and } \gamma_{kr} < 0 \\ w_k + \gamma_{kr} & \text{if } r\epsilon N_s \text{ and } \gamma_{kr} \geq 0 \\ \infty & \text{otherwise.} \end{cases}$$

These are the amounts by which b_k $(k\epsilon M')$ must be increased or $a_{kr}(k\epsilon M')$ must be decreased before the surrogate ceases to fathom s.

C2) $b_k^+(s) = \max\{w_k, -1-\gamma_k(s)\}$; and

$$a_{kr}^{-}(s) = \begin{cases} b_k^{+}(s) & \text{if } r \epsilon J_s^{!} \\ b_k^{+}(s) & \text{if } r \epsilon N_s \text{ and } y_{kr} < 0 \\ \max\{w_k + y_{kr}, -1 - \gamma_k(s) + a_{kr}\delta_{kr}\} & \text{if } r \epsilon N_s \text{ and } y_{kr} \geq 0; \\ & \text{where } \delta_{kr} = \begin{cases} 1 & \text{if } a_{kr} \geq 0, \\ 0 & \text{otherwise.} \end{cases} \\ \infty & \text{otherwise.} \end{cases}$$

hese are the amounts by which $b_k (k \epsilon M')$ must be increased or $a_{kr} (k \epsilon M')$ must be decreased before ɔth the surrogate and constraint k cease to fathom s.

Notice that this approach requires knowledge of the weights u_k for $k \epsilon M$ as well as $\gamma_{m+1}(s)$. Since ıe u_k are dependent upon s, they can be stored or recomputed. Storage of the weights would seem to lace intolerable demands upon system memory. If the weights are easily computed, however, (such as =1 if $t_i(s) > 0$, and 0 otherwise), then the approach can be efficiently implemented. If, on the other ınd, the weights are obtained at some expense, such as from the dual variables associated with the ɔtimal solution to an imbedded linear program, then their recomputation becomes unattractive. In ıis case, some type of compact storage arrangement must be resorted to. (If, for example, weights are mited to one of eight values, then only three bits ($=\log_2 8$) would be required for each weight u_k, for ϵM.)

An alternative approach (more attractive from a computational viewpoint) comes at the cost of sing the refinements on $b_k^{+}(s)$ and $a_{kr}^{-}(s)$ indicated above. When s is fathomed by the surrogate, it is laced into $\Lambda_k (k \epsilon M')$ with $\gamma_k(s)$ as defined by (9). For such s, the ranging variables $b_s^{+}(s)$ and $a_{kr}^{-}(s)$ ıould be set to 0. This modification requires that s be examined for any increase in b_k or decrease a_{kr}. Algorithms 2 and 5 would be performed, without change, for postoptimization. This alternative voids the need to store or recompute the weights, u_k, associated with each partial solution fathomed y the surrogate as described in cases C1 and C2.

In addition to surrogate constraints, some implicit enumeration algorithms (5) solve an imbedded ıear program (ILP) in the free variables. In this case, a partial solution is fathomed if one of the follow-ıg three conditions prevails.

C1) ILP is infeasible.

C2) ILP's optimal objective function value is greater than the best integer solution obtained to date.

C3) ILP's optimal solution satisfies the integrality constraint; and thus becomes the new best solution obtained to date.

Let θ_i, $(i=1, 2, 3)$ be the set of partial solutions fathomed by ILP due to condition i. Figure 1 ıdicates the sets whose partial solutions must be examined for the various parameter changes in-ɔlved in postoptimization. Actually, changes within certain ranges may be made without examining ıy partial solutions. These ranges can be obtained by applying standard linear programming post-ɔtimization techniques to ILP.

The indicated methods for accommodating surrogate constraints and imbedded linear programs re not without cost. By deferring the identification of individual fathoming constraints to the post-ɔtimality phase, the quality of the ranging information provided is significantly reduced. Moreover, single parameter change will, in all likelihood, result in more partial solutions being subjected to fur-ıer examination. The exact extent of these deleterious effects is unknown. A more precise statement ın the utility of surrogate constraints is thus deferred until empirical evidence has been obtained.

Changes being considered during postoptimization	Sets whose partial solutions must be examined
Decrease a_{ij} or increase b_i.....	$\theta_1, \theta_2, \theta_3$
Decrease c_j.....	θ_2, θ_3
All other changes.....	θ_3

FIGURE 1. *Partial solutions fathomed by ILP whose completions must be examined during postoptimization.*

IV. NUMERICAL EXAMPLE

The example presented below, which was originally used by Roodman, illustrates the changes and extensions we have made to his procedures. Consider the following problem.

Minimize: $$z = 2x_1 + 5x_2 + 5x_3 + 6x_4 + 4x_5 + x_6 + 8x_7 + x_8$$

Subject to:
$$-2x_1 + 2x_4 - 6x_5 + x_6 - x_7 + 2x_8 \leq -5$$
$$-4x_1 + 11x_2 - 11x_3 - 7x_4 + 4x_5 + 3x_6 - 5x_7 + x_8 \leq -6$$
$$x_2 + x_3 + x_4 - x_5 - 2x_6 + x_8 \leq 0$$
$$x_j = 0 \text{ or } 1, \quad j = 1, \ldots, 8.$$

The sequence of fathomed partial solutions tabulated below was obtained using Algorithm 1 and Balas' generate rule [1]. (That is, generate the variable $j \epsilon N_s$ which maximizes the quantity $\sum_{i \epsilon M} \min \{t_i - a_{ij}, 0\}$.) Underscored variables are fixed at zero, while the others listed are fixed at one. Free variables are not listed.

Partial solution, s	$z(s)$	Fathoming constraint, k	$\gamma_k(s)$
$s_1 = (7, \underline{5})$	8	1	-2
$s_2 = (7, 5, 4)$	18	feasible	
$s_3 = (7, 5, \underline{4}, 2)$	17	2	-1
$s_4 = (7, 5, \underline{4}, \underline{2}, \underline{3})$	12	2	-1
$s_5 = (7, 5, \underline{4}, \underline{2}, 3)$	17	feasible	
$s_6 = (\underline{7}, \underline{5})$	0	1	-3
$s_7 = (\underline{7}, 5, 3)$	9	feasible	
$s_8 = (\underline{7}, 5, \underline{3}, \underline{1})$	4	2	-3
$s_9 = (\underline{7}, 5, \underline{3}, 1, 2)$	11	0, 2	$-2, -10$
$s_{10} = (\underline{7}, 5, \underline{3}, 1, \underline{2}, 4)$	12	0	-3
$s_{11} = (\underline{7}, 5, \underline{3}, 1, \underline{2}, \underline{4})$	6	2	-6

Thus, Λ can be partitioned as follows: $\Lambda_f = \{s_2, s_5, s_7\}$; $\Lambda_0 = \{s_{10}\}$; $\Lambda_1 = \{s_1, s_6\}$; $\Lambda_2 = \{s_3, s_4, s_8, s_{11}\}$; $\Lambda_3 = \emptyset$; and $\Lambda_g = \{s_9\}$.

Values of the entities $b_k^+(s)$, $b_k^-(s)$, $a_{kr}^-(s)$, $a_{kr}^+(s)$, $c_r^-(s)$, and $c_r^+(s)$, (as defined for the relevant $s \in \Lambda$), are tabulated below to assist the reader in following our ranging formulas.

		$a_{kr}^-(s_h)$								$b_k^+(s_h)$
k	h \ r	1	2	3	4	5	6	7	8	
1	1	1	1	1	3	∞	2	1	3	1
	6	2	2	2	4	∞	3	∞	4	2
2	8	∞	13	∞	2	2	5	∞	3	2
	11	5	∞	∞	∞	5	5	∞	6	4
3	Nil	∞	∞	∞	∞	∞	∞	∞	∞	∞

Hence, $(b_k^+) = (1, 2, \infty)$, and

$$(a_{kr}^+) = \begin{pmatrix} 1 & 1 & 1 & 3 & \infty & 2 & 1 & 3 \\ 5 & 13 & \infty & 2 & 2 & 5 & \infty & 3 \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \end{pmatrix}$$

		$a_{kr}^+(s_h)$								$b_k^-(s_h)$
k	h \ r	1	2	3	4	5	6	7	8	
1	7	∞	∞	1	∞	1	∞	∞	∞	1
2	7	∞	∞	1	∞	1	∞	∞	∞	1
3	7	∞	∞	0	∞	0	∞	∞	∞	0

Hence, $(b_k^-) = (1, 1, 0)$, and

$$(a_{kr}^-) = \begin{pmatrix} \infty & \infty & 1 & \infty & 1 & \infty & \infty & \infty \\ \infty & \infty & 1 & \infty & 1 & \infty & \infty & \infty \\ \infty & \infty & 0 & \infty & 0 & \infty & \infty & \infty \end{pmatrix}$$

$c_r^-(s_h)$								
h \ r	1	2	3	4	5	6	7	8
2	2	5	5	6	4	1	8	1
5	2	5	5	6	4	1	8	1
10	2	5	5	3	4	1	8	1

Hence, $(c_r^-) = (2, 5, 5, 3, 4, 1, 8, 1)$.

$c_r^+(s_h)$								
h \ r	1	2	3	4	5	6	7	8
2	∞	∞	9	∞	∞	∞	∞	∞
5	∞	∞	∞	∞	∞	∞	∞	∞
10	∞	∞	3	∞	∞	∞	∞	∞

Hence, $(c_r^+) = (\infty, \infty, 3, \infty, \infty, \infty, \infty, \infty)$.

In comparing our results with Roodman's, one observes that our procedure indicates that c_2 may be decreased by at least five units ($c_2^- = 5$), whereas Roodman's yields two units. This difference is the direct result of s_9 being omitted from our calculation of c_2^-, since it is fathomed by constraint 0, as well as constraint 2. The other differences between our numerical values and Roodman's result from our consistent use of closed intervals when defining the ranging formulas.

REFERENCES

[1] Balas, E., "An Additive Algorithm for Solving Linear Programs with Zero-One Variables," Operations Res. *13*, 517–546 (1965).

[2] Balas, E., "Discrete Programming by the Filter Method," Operations Res. *15*, 915–957 (1967).

[3] Garfinkel, R., and G. L. Nemhauser, *Integer Programming* (John Wiley and Sons, New York, N.Y., 1970).

[4] Geoffrion, A. M., "Integer Programming by Implicit Enumeration and Balas Method," SIAM Review *9*, 178–190 (1967).

[5] Geoffrion, A. M., "An Improved Implicit Enumeration Approach for Integer Programming," Operations Res. *13*, 879–919 (1968).

[6] Glover, F., "A Multiphase-Dual Algorithm for the Zero-One Integer Programming Problem," Operations Res. *13*, 879–919 (1968).

[7] Glover, F., "Surrogate Constraints," Operations Res. *16*, 741–749 (1968).

[8] Knuth, D. E., *The Art of Computer Programming Volume 1/Fundamental Algorithms* (Addison-Wesley, Reading, Mass., 1969).

[9] Roodman, G., "Postoptimality Analysis in Zero-One Programming by Implicit Enumeration," Nav. Res. Log. Quart. *19*, 435–447 (1972).

A SPECIAL CASE OF THE $3 \times n$ FLOW SHOP PROBLEM *

Fennell Burns

TRW Systems Group
Huntsville, Alabama

John Rooker

Science Applications, Inc.
Huntsville, Alabama

Abstract

Johnson [2] in 1954 solved the two machine flow shop problem by giving an argument for a sufficient condition of optimality and by stating an efficient algorithm which produces a solution via satisfaction of the sufficient condition. Moreover, Johnson solved two special cases of the corresponding three machine flow shop problem. Since that time, six other special cases have been solved, two contributed by Arthanari and Mukhopadhyay [1], two by Smith, Panwalkar, and Dudek [3], and two of a different nature by Szwarc [5]. This paper contributes an extension to one of the classes described by Szwarc.

1. INTRODUCTION

The three machine flow shop problem is mathematically stated as follows (see e.g., Johnson [2] or Szwarc [5]):

Given a positive integer $n > 1$ and $3n$ non-negative numbers

$$A_i, B_i, C_i, i = 1, 2, \ldots, n \tag{1}$$

determine a permutation $p = (p_1, p_2, \ldots, p_n)$ of the integers $1, 2, \ldots, n$ such that the function f defined by

$$f(p) = \max_{1 \le u \le v \le n} [K_u(p) + H_v(p)] \tag{2}$$

is minimized, where

$$K_u(p) = \sum_{i=1}^{u} A_{p_i} - \sum_{i=1}^{u-1} B_{p_i} \tag{3}$$

and

$$H_v(p) = \sum_{i=1}^{v} B_{p_i} - \sum_{i=1}^{v-1} C_{p_i}. \tag{4}$$

* This work was performed partially in support of Contract DAHC60-71-C-0049 of the Ballistic Missile Defense Advanced Technology Center.

The numbers A_i, B_i, C_i in (1) represent for job i the processing times on machine A, B, C respectively. We shall refer to these numbers as the jobset. For any permutation p (also called a schedule) the value $f(p)$ in (2) is the total idle time of machine C after all n jobs have been processed. It is assumed, here, that the processing order for each job is machine A, B, C.

The authors' interest in this problem derived from the examination of real-time software requirements in ballistic missile defense activities involving scheduling of computer and radar resources. Due to the fact that time response requirements are critical in this environment efficient scheduling algorithms are of premium importance.

All of the known special cases, including the one we give here, pivot on a fundamental theorem proved by Smith [4].

THEOREM 1: A sufficient condition that $f(\bar{p}) \leq f(p)$ for all permutations p is that:

(i) there is a real-valued function g of ordered pairs of elements such that if p is any permutation and p' a permutation obtained from p by the interchange, for some j, of the jth and $(j+1)$ st elements, then $f(p) \leq f(p')$ if $g(p_j, p_{j+1}) \leq g(p_{j+1}, p_j)$ and

(ii) $\bar{p}$ is such that the ith element precedes the jth element if $g(i, j) < g(j, i)$.

2. SPECIAL CASE

Consider a jobset which satisfies the following condition:

(S) $$[\min(A_i, B_j) - \min(A_j, B_i)]\,[\min(B_i, C_j) - \min(B_j, C_i)] \geq 0 \qquad i, j = 1, 2, \ldots, n.$$

The implication, here, is that for any i, j from 1 to n the inequalities

(5) $$\min(A_i, B_j) \leq \min(A_j, B_i)$$

and

(6) $$\min(B_i, C_j) \leq \min(B_j, C_i)$$

must hold or the inequalities with "$\leq$" in (5) and (6) replaced by "$\geq$".

EXAMPLE: Let a jobset be defined by

i	A_i	B_i	C_i
1	4	1	1
2	6	10	8
3	5	12	10

Here, (S) is satisfied but none of the other special cases apply.

3. DETERMINATION OF $g(i, j)$

THEOREM 2: Suppose the jobset satisfies (S). Then g defined by

(7) $$g(i, j) = \min(A_i + B_i, B_j + C_j, A_i + C_j)$$

satisfies condition (i) of Theorem 1.

PROOF: Let $p=(p_1, p_2, \ldots, p_n)$ be any permutation, j an integer such that $1 \leq j \leq n$, and p' the permutation obtained from p by interchanging the jth and $(j+1)$st elements. Suppose that

$$g(p_j, p_{j+1}) \leq g(p_{j+1}, p_j) \tag{8}$$

where g is defined in (7). It is necessary to show that $f(p) \leq f(p')$.

The function f defined by (2) can be written in the equivalent form

$$f(p) = \max_{1 \leq i \leq 6} X_i(p) \tag{9}$$

where

$$X_1(p) = \max_{1 \leq u \leq v \leq j-1} [K_u(p) + H_v(p)]$$

$$X_2(p) = \max_{1 \leq u \leq j-1} K_u(p) + \max \{H_j(p), H_{j+1}(p)\}$$

$$X_3(p) = \max_{1 \leq u \leq j-1} K_u(p) + \max_{j+2 \leq v \leq n} H_v(p)$$

$$X_4(p) = \max \{K_j(p) + H_j(p), K_{j+1}(p) + H_{j+1}(p), K_j(p) + H_{j+1}(p)\}$$

$$X_5(p) = \max \{K_j(p), K_{j+1}(p)\} + \max_{j+2 \leq v \leq n} H_v(p)$$

$$X_6(p) \max_{j+2 \leq u \leq v \leq n} [K_u(p) + H_v(p)].$$

Due to the fact that

$$K_u(p) = K_u(p') \qquad u \neq j \text{ or } j+1$$

and

$$H_v(p) = H_v(p') \qquad v \neq j \text{ or } j+1$$

it follows that $X_i(p) = X_i(p')$, $i = 1$, 3, and 6. It is straight forward to show that (8) is equivalent to $X_4(p) \leq X_4(p')$. Now, we will show that

$$X_2(p) \leq X_2(p') \tag{10}$$

and

$$X_5(p) \leq X_5(p') \tag{11}$$

from which it follows that $f(p) \leq f(p')$.

First note the following simple property. Suppose u, v, w are numbers.

(T) If $u \leq v$ then $\min(u, w) \leq \min(v, w)$.

Next we prove:

LEMMA 1: Let g be defined by (7).

(i) If $g(i, j) \leq g(j, i)$ then

$$\min(A_i, B_j) + \min(B_i, C_j) \leq \min(A_j, B_i) + \min(B_j, C_i) \tag{12}$$

(ii) If

$$\min(A_i, B_j) + \min(B_i, C_j) < \min(A_j, B_i) + \min(B_j, C_i) \tag{13}$$

then $g(i, j) < g(j, i)$.

PROOF: Apply (T) with $u = g(i, j)$, $v = g(j, i)$, *and* $w = B_i + B_j$. Then $\min(u, w) \leq \min(v, w)$ is the same as

$$\min(A_i+B_i, B_j+C_j, A_i+C_j, B_i+B_j) \leq \min(A_j+B_j, B_i+C_i, A_j+C_i, B_i+B_j)$$

which is equivalent to (12). Therefore (i) is true. Part (ii) is just the contrapositive of (i) with i and j interchanged.

Now, return to the theorem's proof, where we have assumed that (8) holds. From Lemma 1 inequality (12),

$$\min(A_{p_j}, B_{p_{j+1}}) + \min(B_{p_j}, C_{p_{j+1}}) \leq \min(A_{p_{j+1}}, B_{p_j}) + \min(B_{p_{j+1}}, C_{p_j}). \tag{14}$$

At this point, we envoke the fact that the jobset satisfies (S). So that (14), in view of the implication involving inequalities (5) and (6), implies

$$\min(A_{p_j}, B_{p_{j+1}}) \leq \min(A_{p_{j+1}}, B_{p_j}) \tag{15}$$

and

$$\min(B_{p_j}, C_{p_{j+1}}) \leq \min(B_{p_{j+1}}, C_{p_j}). \tag{16}$$

But inequalities (15) and (16) are equivalent to those in (11) and (10) respectively. Thus, the theorem is proved.

4. SPECIFICATION OF AN ALGORITHM WHEN THE JOBSET SATISFIES (S)

The solution procedure described below is similar to Johnson's two-stage algorithm in that one works from both ends toward the middle.

ALGORITHM:

1. Let $N = \{1, 2, \ldots, n\}$ and construct the permutation $p = (p_1, \ldots, p_n)$ as indicated.
2. Determine integers q, r, s, t in N such that

   (i) $A_q + B_q = \min_{i \in N} A_i + B_i$,

   (ii) $B_r + C_r = \min_{i \in N} B_i + C_i$,

   (iii) $A_s + C_t = \min_{\substack{i, j \in N \\ i \neq j}} A_i + C_j$.

3. If $A_q+B_q \leq \min(B_r+C_r, A_s+C_t)$ then let $p_1=q$.
4. If $B_r+C_r < A_q+B_q$ and $B_r+C_r \leq A_s+C_t$ then let $p_n=r$.
5. If $A_s+C_t < \min(A_q+B_q, B_r+C_r)$ then one of the following conditions is true:
   (i) $A_s \leq C_t$
   (ii) $C_t < A_s$

   If (i) holds, let $M=\{j|A_s=A_j, j\epsilon N, j\neq t\}$ and denote the number of integers in M by l. Then let $p_1, p_2, \ldots, p_l$ be determined by applying Johnson's two-stage rule to the two-stage jobset α_i, β_i, where $\alpha_i=A_i+B_i$, $\beta_i=\min(A_i, B_i)+C_i$ with $i\epsilon M$.

   If (ii) holds, let $M=\{j|C_t=C_j, j\epsilon N, j=s\}$ and denote the number of integers in M by l. Then let $p_{n-l+1}, \ldots, p_n$ be determined by applying Johnson's two-stage rule to the two-stage jobset α_i, β_i, where $\alpha_i=A_i+\min(B_i, C_i)$, $\beta_i=B_i+C_i$ with $i\epsilon M$.
6. Delete those elements scheduled from N.
7. Repeat steps 2 through 6 until all elements of N are scheduled.

To establish the optimality of the schedule p generated by the algorithm we need:

THEOREM 3: Suppose the jobset satisfies (S). Then the algorithm produces a permutation p that satisfies condition (ii) of Theorem 1 with g defined by (7).

PROOF: Let p be the permutation constructed by the algorithm. It is necessary to show that for each pair of integers i, j with $1 \leq i \leq j \leq n$ that $g(p_j, p_j) \leq g(p_j, p_j)$ is satisfied. Due to the recursive nature it is sufficient to show that the first iteration determines p_1 or p_n or $p_1, p_2, \ldots, p_l$ or $p_{n-l+1}, \ldots, p_n$ correctly.

Suppose on the first iteration, the inequality in step 3 is satisfied. In view of conditions (i), (ii), and (iii) in 2, for any integer $j\epsilon N$, it is clear that $g(p_1, j) \leq g(j, p_1)$.

Suppose the inequalities in step 4 are satisfied. Then it is easy to see that the conditions in step 2 imply for any $j\epsilon N$, $g(j, p_n) \leq g(p_n, j)$.

Suppose the inequality in step 5 holds. In the event that condition (i) is true then for $p_i \epsilon M$, $A_{p_i} < B_{p_i}$ and $A_{p_i} < A_j, j \neq t$ or $j \notin M$, where $M=\{p_1, \ldots, p_l\}$. Therefore:

$$\min(A_{p_i}, B_j) < \min(A_j, B_{p_i}) \qquad j \neq t \text{ or } j \notin M.$$

But the jobset satisfies (S), thus

$$\min(B_{p_i}, C_j) \leq \min(B_j, C_{p_i}) \qquad j \neq t \text{ or } j \notin M.$$

Adding these last two inequalities and using (13) of Lemma 1 proves that:

$$g(p_i, j) \leq g(j, p_i) \qquad j \neq t \text{ or } j \notin M.$$

In case $j=t$, from condition (iii) of step 2 and the first inequality of step 5, it follows that $g(p_i, t) < g(t, p_i)$. If $i, j\epsilon M$ then $A_i=A_j$. This means that $g(i, j) \leq g(j, i)$ is equivalent to

$$\min(A_i+B_i, \min(A_j, B_j)+C_j) \leq \min(A_j+B_j, \min(A_i, B_i)+C_i).$$

But this inequality is satisfied by the schedule produced from Johnson's two-stage algorithm applied

as described in step 5. Therefore, $g(p_i, p_j) \leq g(p_j, p_i) 1 \leq i \leq j \leq l$. The case for condition (ii) is similar.

Therefore, the algorithm schedules p_1 or p_n or $p_1, \ldots, p_l$ or $p_{n-l+1}, \ldots, p_n$ correctly on the first iteration and the theorem follows.

If a jobset satisfies (S) and a schedule p is constructed using the algorithm, then Theorem 1, 2, and 3 collectively imply that p is optimal.

EXAMPLE. Consider the jobset

i	1	2	3	4	5	6	7	8
A_i	1	7	7	4	1	6	17	17
B_1	10	7	6	5	9	6	15	15
C_i	12	3	3	5	11	2	16	17

Condition (S) is satisfied and the algorithm requires six iterations producing in order: $p_1=5$, $p_2=1$ by step 5(i); $p_8=6$ by step 5(ii); $p_6=2$, $p_7=3$ by step 5(ii); $p_3=4$ by step 3; $p_5=7$ by step 4; $p_4=8$ by step 3. Thus $p=(5, 1, 4, 8, 7, 2, 3, 6)$ is an optimal permutation.

5. ADDITIONAL REMARKS

(a) Not only will the algorithm determine an optimal sequence for jobsets which satisfy (S) but also the two cases discovered by Johnson.

(b) Theorem 2 as stated applies only to the class of jobsets which satisfy condition (S). But it is clear on examining the proof, the result applies more generally. The following corollary focuses attention on this matter. Let the functions f and g be those defined in (2) and (7) respectively.

COROLLARY: Suppose j is an integer, $1 \leq j < n$, and the jobset and schedule $p=(p_1, \ldots, p_j, p_{j+1}, \ldots, p_n)$ have the property that:

(i) the inequality in (S) is satisfied for the integers p_j and p_{j+1} and

(ii) $g(p_j, p_{j+1}) \leq g(p_{j+1}, p_j)$.

Then $f(p) \leq f(p')$ where p' is obtained from p by interchanging the jth and $(j+1)$st elements.

This follows immediately from Theorem 2 since the hypothesis here is the only hypothesis used in proving Theorem 2.

The value to be added, here, is the set of admissible permutations may be reduced in number. Consider a jobset which has two jobs i and j that satisfy (i) and (ii) of the corollary. Then any permutation where i is adjacent to and follows j may be eliminated from consideration. Evidently, there are $(n-1)!$ such permutations.

(c) One more item of interest is the following situation. If there are only two jobs in the jobset then inequality (8) with g defined in (7) determines the optimal ordering, regardless of any special conditions. This is true because (8) is equivalent to $X_4(p) \leq X_4(p')$ from (9) and the terms in $X_4(p)$ are the only ones to consider in the two job case.

REFERENCES

[1] Arthanari, T. S., and A. C. Mukhopadhyay, "A Note on a Paper by W. Szwarc," Nav. Res. Log. Quart., *18*, 135–138 (1971).

[2] Johnson, S. M., "Optimal Two and Three Stage Production Schedules With Set-Up Time Included," Nav. Res. Log. Quart., *1*, 61–68 (1954).

[3] Smith, M. L., S. S. Panwalkar, and R. A. Dudek, "Flowshop Sequencing Problem With Ordered Processing Time Matrices," Management Science, *21*, 554–549 (1975).

[4] Smith, Wayne E., "Various Optimizers for Single-Stage Production," Nav. Res. Log. Quart., *3*, 59–66 (1956).

[5] Szwarc, Wlodzimierz, "Mathematical Aspects of the $3 \times n$ Job-Shop Sequencing Problem," Nav. Res. Log. Quart., *21*, 145–153 (1974).

A NOTE ON "FAILURES-BEFORE-FIRST-SUCCESS RUN SIZE" IN LIFE TESTING

Prem K. Goel

Purdue University

SUMMARY

We define "Failures-Before-First-Success Run Size" $Z(x)$ as the number of failures (deaths prior to the completion of the mission) before first success (one member survives past the duration, x, of the mission) in a sequence of i.i.d. trials with probability of success $(1-F(x))$, where $F(x)$ is the life distribution. If F belongs to a class of distributions called NBU (New Better than Used), then the following inequalities are established.

If $g(x)$ denotes the expected value of $Z(x)$, then

(1) $$g(x+y) \geq g(x)g(y) \qquad \text{for all nonnegative } x, y$$

If $\nu(x)$ denotes the variance of $Z(x)$, then

(2) $$\nu(x+y) \geq \nu(x)\nu(y) \qquad \text{for all nonnegative } x, y$$

A simple application of (1) and (2) is also suggested.

INTRODUCTION

Let $F(x)$ denote the life distribution of a component. We assume that $F(x)=0$ for $x<0$. A certain class of life distributions that arise naturally and play a crucial role in the comparison and study of some maintenance policies is defined as below.

DEFINITION

A life distribution F is said to be New Better than Used (NBU) (Marshall and Proschan [4]) if:

(3) $$1-F(x+y) \leq (1-F(x))(1-F(y)) \qquad \text{for all } x, y \geq 0.$$

This means that a used item has stochastically smaller remaining life than does a new item, or mathematically it means that $-\log(1-F(t))$ is superadditive. Marshall and Proschan [5] show that NBU distributions arise naturally in replacement policies. They also point out that NBU contains the class of IFRA (Increasing Failure Rate Average) and IFR (Increasing Failure Rate) distributions.

Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. Bernoulli trails with probability of success $(1-F(x))$, and let $Z(x)$ be the number of failures in the above sequence before the first success. It is well known that $Z(x)$ has a geometric distribution with probability of success $(1-F(x))$, i.e.,

$$P\{Z(x)=k\}=(1-F(x))[F(x)]^k \qquad k=0, 1, 2, \ldots,$$

and

$$E[Z(x)]=\{F(x)/(1-F(x))\}=g(x) \text{ (say)} \tag{4}$$

and

$$\mathrm{Var}[Z(x)]=\{F(x)/(1-F(x))^2\}=\nu(x) \text{ (say).} \tag{5}$$

We shall call $Z(x)$ the "Failures-Before-First-Success Run Size." The inequality (1) is an extension of an unsolved problem proposed by Bloom [1, problem #5942, p. 1147]. Bloom proposed it for the normal distribution and its proof is given by Neuts [6] and Clarke et al. [2]. The proof of inequalities (1) and (2) is not too difficult for NBU distributions. However, a contribution of this note is to give a probabilistic interpretation of $F(x)/(1-F(x))$ in life testing problems and its applications.

PROPOSITION 1: Let $g(x)$ denote the "Expected Failures-Before-First-Success Run Size" for a life time distribution $F(x)$ satisfying (3), then $g(x)$ satisfies (1)i.e., $\log g(x)$ is superadditive.

PROOF: Let us rewrite (3) as

$$F(x+y) \geq F(x)+F(y)-F(x)F(y), \tag{6}$$

and use the fact $F(x) \geq F^2(x)$ for all x, we get

$$F(x)+F(y) \geq F^2(x)+F^2(y). \tag{7}$$

Now it follows from (6), (7) and the fact $(F(x)-F(y))^2 \geq 0$, that

$$F(x)F(y) \leq F(x+y). \tag{8}$$

On multiplying (3) and (8), it follows from (4) that (1) holds.

COROLLARY: If F satisfies (3) and $\nu(x)$ denotes the variance of "Failures-Before-First-Success Run Size" then $\nu(x)$ satisfies (2), i.e., $\log \nu(x)$ is superadditive.

PROOF: It follows from (5) that $\nu(x)=g(x)/(1-F(x))$. This fact and division of (1) by (3) imply that (2) holds.

REMARK 1: The restriction $F(x)=0$ for all $x<0$ is needed, since (1) does not hold even for all Symmetric Polya type II distributions (Karlin [3, p. 152] establishes that $(1-F(x))$ is Polya type II if and only if $F(x)$ has IFR). A counter example is given by the double exponential distribution, for which the inequality (1) is reversed in direction. However, for the logistic distribution, equality holds.

REMARK 2: It is observed that if a random variable X, with cdf $F(x)$, is symmetric around zero and that $(1-F(x))$ is Polya type II than $|X|$ is a nonnegative random variable with IFR. Therefore inequality (1) applies to the distribution of $|X|$. Hence it follows from (1) that

$$\frac{P\{|X| \leq x+y\}}{P\{|X|>x+y\}} \geq \frac{P\{|X| \leq x\}}{P\{|X|>x\}} \cdot \frac{P\{|X| \leq y\}}{P\{|X|<y\}}$$

holds.

APPLICATION

In the following discussion, it is assumed that F satisfies (3). Some properties of the random variable $N(x)$, number of failures in $[0, x]$ if replacements are made only upon failure, are discussed in [5]. However, in several situations, e.g., unmanned space explorations, it may not be possible to replace some component immediately upon failure, but it may be necessary that this component works continuously for a successful completion of the mission. Therefore one needs to provide an "adequate" number, $n(x)$, of components in parallel. The mission will not fail as long as one of these survives for the duration of the mission. In such a situation the random variable $Z(x)$, failures-before-first-success run size, is the appropriate choice. If the function $F(x)$ is unknown and $n(x)$ is chosen to be some function of $g(x)=E\{z(x)\}$ and $\nu(x)=\mathrm{Var}\{z(x)\}$, then one needs to estimate $g(x)$ and $\nu(x)$ for a given duration x. However, if the duration, x, is large then it may not be possible to carry out an experiment to estimate $g(x)$ and $\nu(x)$. Therefore an alternative is to estimate $g(t)$ and $\nu(t)$ from a shorter duration experiment (say, e.g., $t=x/2$ or $x/4$) and then obtain estimated lower bounds of $g(x)$ and $\nu(x)$ by using inequalities (1) and (2).

If the "cost" pertaining to this component is proportional to $n(x)$, then the inequalities (1) and (2) give some idea about the behavior of this "cost" as a function of the mission's duration.

ACKNOWLEDGMENT

I would like to thank the referee for some helpful comments and suggestions.

REFERENCES

[1] Bloom, D. M., "Unsolved Problem #5942," Amer. Math. Monthly *80*, 1147 (1973).

[2] Clarke, L. E., "Solution to Problem #5942," Amer. Math. Monthly *82*, 186–187 (1975).

[3] Karlin, S., *Total Positivity* (Stanford Univ. Press, Stanford, Calif., 1968), Vol. I.

[4] Marshall, A. W. and F. Proschan, "Maximum likelihood estimation for distributions with Monotone failure rate," Ann. Math. Stat. *36*, 69–71 (1965).

[5] Marshall, A. W. and F. Proschan (1970). Classes of Distributions Applicable in Replacement, With Renewal Theory Implications. Proc. Sixth Berk. Sym. Math. Stat. Prob., I (1970), pp. 395–415.

[6] Neuts, M. F., "Solution to Bloom's Problem," private communication (1974).

A NOTE ON THE OPTIMAL SCHEDULING OF TWO PARALLEL PROCESSORS

Victor Zaloom

Auburn University

and

David Vatz

TRW/Systems Group
McLean, Virginia

ABSTRACT

The problem of minimizing mean flow time of two parallel processors is discussed. Prior results are briefly reviewed. A dynamic programming algorithm is presented which minimizes mean flow time for a set of n preordered jobs on two nonequivalent parallel processors. The algorithm is illustrated with an example problem. The computational experience is presented which illustrates the efficiency of the algorithm.

SUMMARY

A service facility containing two parallel processors is available to process jobs. The objective is to minimize mean flow time. This objective is accomplished by optimally partitioning and sequencing the jobs to the two processors. For the case in which no order relationship exists among the jobs and the processors have equal capabilities, it is known (see Conway, Maxwell, and Miller [1]) that mean flow time is minimized by ordering the jobs according to the shortest processing time. The jobs are then sequenced to the processors in alternating fashion (i.e., jobs 1, 3, 5, . . . go to processor 1 and jobs 2, 4, 6, . . . go to processor 2).

If there exists a predefined order on the jobs, such as the order of arrival, or of importance, which must be preserved on each processor, (see figure 1), then a dynamic programming algorithm presented by Mehta, Chandrasekaran, and Emmons [2] is capable of minimizing mean flow time for two equivalent parallel processors. The algorithm of Mehta, et al., was extended by Rothkopf [3] (see this issue of NRLQ) for the case in which the two parallel processors are not necessarily equivalent.

In an independent effort, the authors have also extended the results of Mehta, Chandrasekaran, and Emmons. The purposes of this note are to illustrate the algorithm with an example problem, relate some of the author's computational experience concerning the algorithm and discuss a number of interesting applications and extensions.

ROTHKOPF'S MODEL

Rothkopf's basic dynamic programming relationships are

$$f_i(k) = (a+k)T_{i,1} + B_{i,1} + \min\{f_{i-1}(k-1), g_i - 1(i-k)\} \tag{1}$$

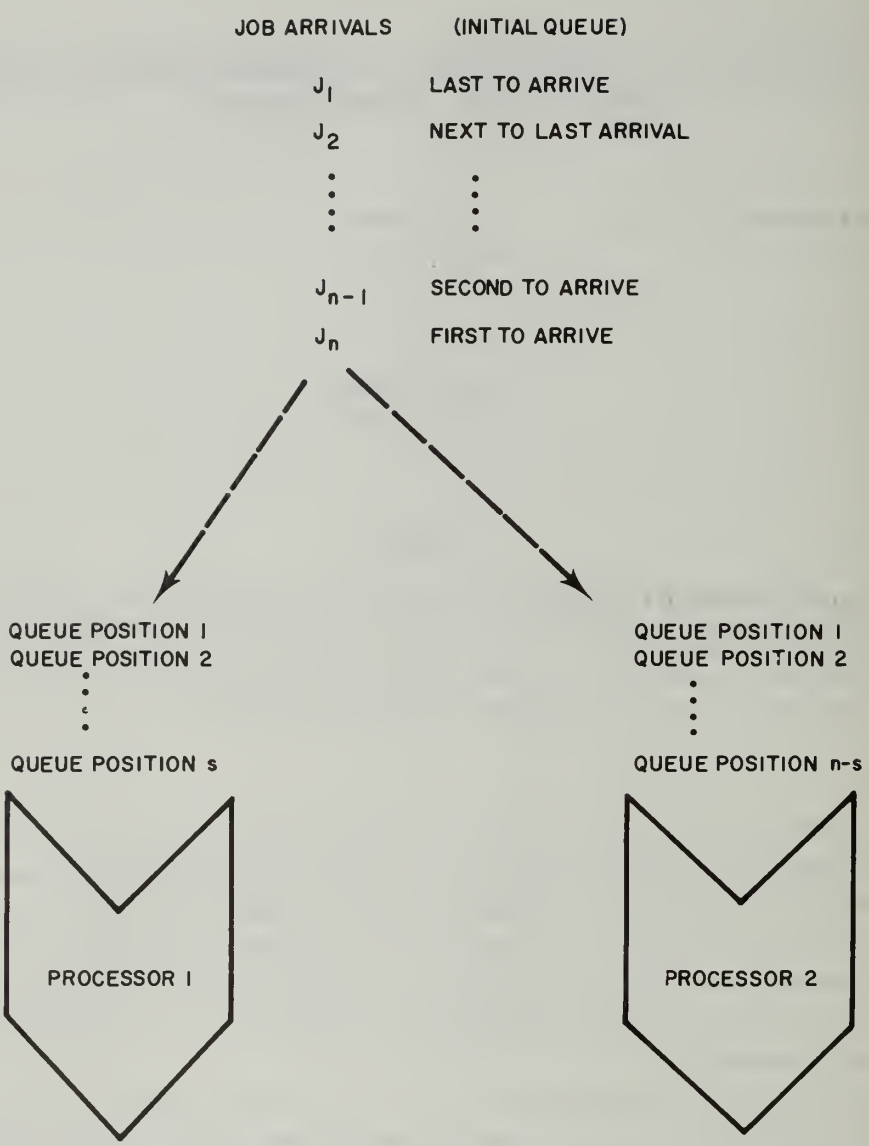

FIGURE 1. The job assignment process.

and

$$g_i(k) = (b+k)T_{i,2} + B_{i,2} + \min\{f_{i-1}(i-k),\ g_i - 1(k-1)\}, \tag{2}$$

where $f_i(k)$ is the minimum total cost for completing only jobs $1, 2, \ldots, i$ with job i in position k from the end on processor 1; $g_i(k)$ is a similar relationship for processor 2. The constants a and b represent costs of time on processors 1 and 2, respectively. The term $kT_{i,j}$ represents the additional flow time due to job i for the $k-1$ jobs which follow job i on processor j plus the flow time of job i. $B_{i,j}$ is a special cost incurred if job i is scheduled on processor j. For our formulation $B_{i,j}=0$ for $i=1, 2, \ldots, n$ and $j=1,2$. Rothkopf defines two artificial tasks, $n+1$ and $n+2$ such that $B_{n+1,2}=B_{n+2,1}=\infty$ and $B_{n+1,1}=B_{n+2,2}=0$. The authors' formulation has no need for these artificial tasks since no special costs are

assumed. Further definitions in Rothkopf's formulation are:

$f_1(1) = (a+1)T_{1,1} + B_{1,1}$, $g_1(1) = (b+1)T_{1,2} + B_{1,2}$ and $f_i(0) = f(i+1) = g_i(0) = g_i(i+1) = \infty$, $i = 1, 2, \ldots, n$.

With the above definitions and the further assumption that the constants a and b are zero, Equations (1) and (2) can be rewritten as (3) and (4), respectively:

(3) $$f_i(k) = kT_{i,1} + \min \{f_{i-1}(k-1), g_{i-1}(i-k)\}$$

and

(4) $$g_i(k) = kT_{i,2} + \min \{f_{i-1}(i-k), g_{i-1}(k-1)\}.$$

AN EXAMPLE

This example problem will be solved in detail to illustrate the algorithm. Table 1 contains the matrix of processing times, $T_{i,j}$ for each job on each processor.

TABLE 1. *Example Problem—Processing Times*

Processors ↓	Jobs			
	J_4	J_3	J_2	J_1
1	4	9	7	10
2	6	3	10	15

Stage 1:

$f_1(1) = 1T_{1,1} = 10$

$g_1(1) = 1T_{1,2} = 15$

Stage 2:

$f_2(1) = 1T_{2,1} + \min \{f_1(0), g_1(1)\} = 7 + \min \{\infty, 15\} = 22$

$f_2(2) = 2T_{2,1} + \min \{f_1(1), g_1(0)\} = 2(7) + \min \{10, \infty\} = 24$

$g_2(1) = 1T_{2,2} + \min \{f_1(1), g_1(0)\} = 10 + \min \{10, \infty\} = 20$

$g_2(2) = 2T_{2,2} + \min \{f_1(0), g_1(1)\} = 20 + \min \{\infty, 15\} = 35$

Stage 3:

$f_3(1) = 1T_{3,1} + \min \{f_2(0), g_2(2)\} = 9 + \min \{\infty, 35\} = 44$

$f_3(2) = 2T_{3,1} + \min \{f_2(1), g_2(1)\} = 18 + \min \{22, 20\} = 38$

$f_3(3) = 3T_{3,1} + \min \{f_2(2), g_2(0)\} = 27 + \min \{24, \infty\} = 51$

$g_3(1) = 1T_{3,2} + \min \{f_2(2), g_2(0)\} = 3 + \min \{24, \infty\} = 27$

$g_3(2) = 2T_{3,2} + \min \{f_2(1), g_2(1)\} = 6 + \min \{22, 20\} = 26$

$g_3(3) = 3T_{3,2} + \min \{f_2(0), g_2(2)\} = 9 + \min \{\infty, 35\} = 44$

Stage 4:

$f_4(1) = 1T_{4,1} + \min\{f_3(0), g_3(3)\} = 4 + \min\{\infty, 44\} = 48$
$f_4(2) = 2T_{4,1} + \min\{f_3(1), g_3(2)\} = 8 + \min\{44, 26\} = 34$
$f_4(3) = 3T_{4,1} + \min\{f_3(2), g_3(1)\} = 12 + \min\{38, 27\} = 39$
$f_4(4) = 4T_{4,1} + \min\{f_3(3), g_3(0)\} = 16 + \min\{51, \infty\} = 67$
$g_4(1) = 1T_{4,2} + \min\{f_3(3), g_3(0)\} = 6 + \min\{51, \infty\} = 57$
$g_4(2) = 2T_{4,2} + \min\{f_3(2), g_3(1)\} = 12 + \min\{38, 27\} = 39$
$g_4(3) = 3T_{4,2} + \min\{f_3(1), g_3(2)\} = 18 + \min\{44, 26\} = 44$
$g_4(4) = 4T_{4,2} + \min\{f_3(0), g_3(3)\} = 24 + \min\{\infty, 44\} = 68$

The backtracking step goes as follows (refer to Figure 2). The minimum total flow time of 34 is achieved by assigning job 4 to processor 1, position 2. Therefore, job 3 must be assigned to position 1 of processor 1 or to position 2 of processor 2. The optimal decision is to assign job 3 to position 2 of processor 2. Likewise, job 2 goes to position 1 of processor 2 and job 1 to position 1 of processor 1.

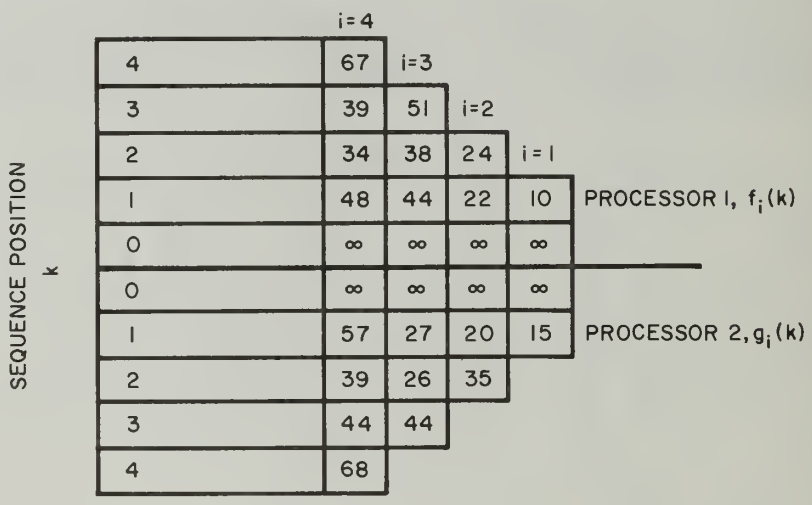

FIGURE 2. Dynamic programming table.

COMPUTATIONAL EXPERIENCE

A Fortran version of the algorithm solved 25 example problems. Each job was assigned an integer time greater than or equal to 1 and less than or equal to 10. No attempt was made to optimize computer time. The problem data appear in Table 2. There were five problems with 15 jobs, five problems with 30 jobs, and so on up to 240 jobs. The computational time to solve each problem is given in seconds on the IBM 370 model 155 computer. The total flow time of the optimal solution is also given in Table 2. The computation times are reasonable even for relatively large numbers of jobs. Also, the small variance of computation times indicates the consistency of the algorithm's performance on different problems with the same number of jobs.

The number of $f_i(k)$ and $g_i(k)$ values to be stored is $n(n+1)$. The number of multiplications, comparisons, and additions totals $3(n)(n+1)$. This is approximately 3 times the computational requirements of the algorithm presented by Mehta et al. [2] for the equivalent processor case. The

algorithm is capable of solving large problems since storage and computational requirements increase on the order of n^2 and since only two stages need to be stored in core at any time. Computational results in Table 2 indicate that as the number of jobs increase from n to $2n$, the computation time appears to be approaching an increase from t to $4t$. In going from 15 to 30, 30 to 60, 60 to 120, and 120 to 240 jobs, average computer time increases by factors of 2.3, 2.6, 3.1, and 3.5, respectively.

TABLE 2. *Computational Data*

Problem Number	Number of jobs	Computation time in seconds	Total flow time
1	15	0.14	259
2	15	0.14	220
3	15	0.15	249
4	15	0.15	179
5	15	0.15	204
6	30	0.34	908
7	30	0.33	767
8	30	0.33	1,057
9	30	0.32	865
10	30	0.33	1,123
11	60	0.87	2,902
12	60	0.87	3,303
13	60	0.87	3,851
14	60	0.87	3,352
15	60	0.87	3,491
16	120	2.75	13,580
17	120	2.81	12,225
18	120	2.65	12,953
19	120	2.66	15,604
20	120	2.66	14,029
21	240	9.32	55,472
22	240	9.37	54,533
23	240	9.34	51,699
24	240	9.40	53,094
25	240	9.26	52,952

REFERENCES

[1] Conway, R., W. Maxwell, and L. Miller, *Theory of Scheduling* (Addison Wesley, Reading Massachusetts, 1967).

[2] Mehta, S., R. Chandrasekaran, and H. Emmons, "Order-Preserving Allocation of Jobs to Two Machines," Nav. Res. Log. Quart. *21*, 361–364 (June 1974).

[3] Rothkopf, Michael H., "A Note On Allocating Jobs to Two Machines," Nav. Res. Log. Quart. *22*, (Dec. 1974).

A NOTE ON ALLOCATING JOBS TO TWO MACHINES

Michael H. Rothkopf

Xerox Palo Alto Research Center
Palo Alto, California

ABSTRACT

A recent article in this journal by Mehta, Chandrasekaran, and Emmons [1] described a dynamic programming algorithm for assigning jobs to two identical parallel processors in a way that minimizes the average delay of these jobs. Their problem has a constraint on the sequence of the jobs such that any group of jobs assigned to a processor must be processed in the order of the sequence. This note has two purposes. First, we wish to point out a relationship between this work and some prior work [2]. Second, we wish to point out that Mehta, Chandrasekaran, and Emmons formulation, slightly generalized, can be used to find the optimum assignment of jobs to two machines in a more general class of problems than they considered including a subclass in which the jobs are *not* constrained to be processed in a given sequence.

STATE OF THE ART

Rothkopf [2] has given a dynamic programming formulation for the problem of scheduling n immediately available jobs on m machines under the following assumptions:

1. Each job must be processed by exactly one machine;
2. Each job, i, has an integer processing time $T_{i,j}$ if it is processed by machine j;
3. Each machine, j, becomes available for processing at a nonnegative integer time t_j^* and is continuously available thereafter;
4. Each job, i, has a monotonic, nondecreasing total waiting cost function, $C_i(t)$;
5. Each job, i incurs a special cost $B_{i,j}$ if it is processed by machine j;
6. There is a cost, $G_0(Y_1, Y_2, \ldots, Y_n)$, associated with a schedule in which machine j completes the processing of the jobs assigned to it at time Y_j; G_0 is a monotonic nondecreasing function of each of its arguments for any set of values for the others; and
7. Each machine must process the jobs assigned to it in reverse numerical order.

Three different conditions under which assumption 7 is nonrestrictive are developed. Rothkopf's dynamic programming formulation is better than complete enumeration. The problem considered by Mehta, Chandrasekaran, and Emmons [1] is a specialization of Rothkopf's in which $m=2$; $T_{i,j}=T_i$; $t_1^*=t_2^*=0$; $C_i(t)=t$, $i=1, 2, \ldots, n$; $B_{i,j}=0$; $i=1, 2, \ldots, n$, $j=1, 2$; and $G_0(Y_1, Y_2)=0$. Their dynamic programming formulation is substantially more efficient than his for this problem although the effort for each formulation grows as the square of the number of jobs.

GENERALIZATION

Consider Rothkopf's problem under the following restrictions: $m=2$; $C_i(t)=t$, $i=1, 2, \ldots, n$; and $G_0(Y_1, Y_2)=aY_1+bY_2$. This problem is considerably more general than Mehta's et al. problem,

but can still be solved by using a slightly modified version of their formulation. This modification preserves the bulk of their formulation's computational advantage.

In order to generalize Mehta's et al. approach to handle the problem considered here, we first define two artificial tasks. These tasks are numbered $n+1$ and $n+2$ and have $T_{n+1,1}=t_1^*$, $T_{n+2,2}=t_2^*$, $B_{n+1,2}=B_{n+2,1}=\infty$, and $B_{n+1,1}=B_{n+2,2}=0$. This takes care of the machine availability assumption (#3 above). Next, we distinguish the two processors in a way that Mehta et al. found unnecessary by defining $f_i(k)$ as the minimum total cost for completing jobs 1, 2, . . ., i (assuming these are the only jobs to be processed), putting job i in position k from the end on processor 1; $g_i(k)$ is defined similarly for processor 2. The basic pair of dynamic programming recurrence relationships are:

$$f_i(k)=(a+k)T_{i,1}+B_{i,1}+\min\{f_{i-1}(k-1),\, g_{i-1}(i-k)\} \tag{1}$$

and

$$g_i(k)=(b+k)T_{i,2}+B_{i,2}+\min\{f_{i-1}(i-k),\, g_{i-1}(k-1)\}, \tag{2}$$

with

$$f_1(1)=(a+1)T_{1,1}+B_{1,1}, \qquad g_1(1)=(b+1)T_{1,2}+B_{1,2}$$

and

$$f_j(0)=f_j(j+1)=g_j(0)=g_j(j+1)=\infty \quad (j=1, 2, \ldots, n+2).$$

Although there may be further simplifications possible in some special cases, the optimal assignment of jobs to machines can always be found by evaluating in succession $f_1(1)$, $g_1(1)$; $f_2(k)$, $g_2(k)$, $k=1, 2$; . . .; $f_{n+2}(k)$, $g_{n+2}(k)$, $k=1, 2, \ldots, n+2$.

AN INTERESTING SPECIAL CASE

If we assume additionally $T_{i,j}=T_i \cdot K_j$, $j=1, 2$, $i=1, 2, \ldots, n$, i.e., the machines have the same relative efficiencies for each job, then we can still use the above algorithm to solve the job scheduling problem when the order-preserving assumption (number 7 above) is relaxed. To do this we number the jobs so that $T_1 \geq T_2 \geq T_3 \geq \ldots, \geq T_n$. Under these conditions it is optimal to process any selection from these n jobs assigned to either machine in reverse numerical order. Therefore, with this numbering we can use assumption 7 without actually constraining the problem.

REFERENCES

[1] Mehta, S., R. Chandrasekaran, and H. Emmons, "Order-Preserving Allocation of Jobs to Two Machines," Nav. Res. Log. Quart. *21*, 361–364 (1974).

[2] Rothkopf, Michael H., "Scheduling Independent Tasks on Parallel Processors," Management Science *12*, 437–447 (1966).

NEWS AND MEMORANDA

List of Referees

The Editors of the Naval Research Logistics Quarterly are grateful to the following individuals for assisting in the review of articles prior to publication.

J. G. Abert
H. Ascher
K. R. Baker
V. Balachandran
E. Balas
J. Balintfy
C. B. Barfoot
R. E. Barlow
M. S. Bazaraa
R. Bechhofer
L. Berkovitz
P. J. Bickel
W. R. Blischke
S. Bonder
V. J. Bowman, Jr.
J. Bracken
P. Bratley
R. Butterworth
A. V. Cabot
A. J. Clark
B. W. Conolly
R. Cottle
R. T. Crow
W. R. Crowston
E. Denardo
W. Dent
C. Derman
R. L. Disney
L. Dolansky
S. D. Dubey
R. A. Dudek
E. J. Dudewicz
E. P. Durbin
P. S. Dwyer
B. Curtis Eaves
R. L. Eisenman
S. E. Elmaghraby
H. Emmons
J. P. Evans
J. Falk
M. J. Fisher
R. L. Francis
M. Geisler
A. Geoffrion
N. D. Glassman
F. Glover
R. Grafton
G. W. Graves
H. J. Greenberg
R. C. Grinold
D. Gross
F. E. Grubbs
W. C. Guenther
J. N. D. Gupta
S. Haber
M. S. A. Hameed
W. W. Hardgrave
C. M. Harris
H. O. Hartley
I. D. Hill
F. S. Hillier
R. G. Hinkle
A. J. Hoffman
W. A. Horn
T. C. Hu
D. L. Iglehart
E. Ignall
S. C. Jaquette
J. J. Jarvis
J. B. Kadane
J. H. K. Kao
S. Kaplan
U. Karmarkar
W. J. Kennedy, Jr.
B. M. Khumawala
L. Kleinrock
P. J. Kolesar
J. Laderman
L. Lasdon
C. E. Lemke
C. E. Lewis
J. G. Lieberman
B. P. Lientz
S. A. Lippman
E. P. Loane
T. J. Lowe
W. F. Lucas
B. J. McDonald
M. J. Magazine
N. R. Mann
P. Marlin
H. F. Martz, Jr.
C. Marshall
K. T. Marshall
C. B. Millham
R. C. Morey
J. A. Muckstadt
F. Murphy
C. Mylander
A. N. Napier, Jr.
M. F. Neuts
I. Olkin
G. Owen
W. R. Pabst, Jr.
J. H. Patterson
W. Pierskalla
E. L. Porteus
N. U. Prabhu
V. J. Presutti, Jr.
F. Proschan
M. R. Rao
H. D. Ratliffe
J. G. Rau
G. Roodman
M. H. Rothkopf
D. P. Rutenberg
T. Saaty
H. M. Salkin
R. Schafer
D. A. Schrady
L. E. Schrage
W. R. Schucany
J. F. Sharp
L. Shaw
M. L. Shooman
M. Shubik
R. Simon
N. D. Singpurwalla
D. Smith
R. Soland
Henry Solomon
Herbert Solomon
W. A. Spivey
V. Srinivasan
D. I. Steinberg
R. E. Steuer
G. Suzuki
W. Szwarc
H. A. Taha
J. G. Taylor
M. E. Thomas
G. Thompson
T. Varley
H. M. Wagner
K. T. Wallenius
A. R. Washburn
A. B. Weiss
L. Weiss
A. C. Sheeler
A. Whinston
T. M. Whitin
J. D. Wiest
W. L. Wilkinson
E. B. Wilson
G. P. Wright
R. D. Young
S. Zacks
V. A. Zaloom
W. J. Zimmer
S. Zionts

INDEX TO VOLUME 22

INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.

NAVAL RESEARCH
LOGISTICS
QUARTERLY

DECEMBER 1975
VOL. 22, NO. 4
NAVSO P-1278

CONTENTS

ARTICLES

OFFICE OF NAVAL RESEARCH
Arlington, Va. 22217